Counter-progagating waves on fluid surfaces and the continuum limit of the Fermi-Pasta-Ulam model

Guido Schneider
Mathematisches Institut
Universität Bayreuth
95440 Bayreuth
Germany
guido.schneider@uni-bayreuth.de

C. Eugene Wayne
Department of Mathematics
Boston University
111 Cummington St.
Boston, MA 02215
USA
cew@math.bu.edu

Abstract

The rigorous approximation of long wavelength motions of fluids by a pair of uncoupled Korteweg-de Vries equations is described. We then extend this result to study the motion of the Fermi-Pasta-Ulam model of coupled nonlinear oscillators.

1 Introduction

Amplitude, or modulation, equations have found a host of uses in applied mathematics. Typically, they are derived via a formal asymptotic analysis, and until recently there were few rigorous results concerning their validity. In particular, in some circumstances, several different equations have been derived as models for the same physical situation, and in the absence of mathematical results justifying the approximations made in their derivations, it is difficult to decide which (if any) of these equations actually provide accurate approximations to the original problem.

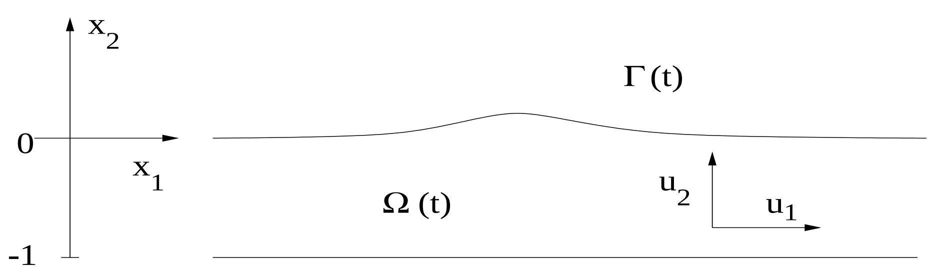
Within the past decade, there has been a considerable advance in the understanding of how one can justify a variety of modulation equations. In particular, the Ginzburg-Landau equation and Nonlinear Schrödinger equation have been shown to provide accurate models in a number of different circumstances [3], [15], [14], [10]. One case that has received particular attention is the problem of counter-propagating waves. This involves systems which support localized waves moving both to the left and the right, and in the physics literature there were several different proposals for the form of the interaction term in the modulation equations for these problems. In both [9] and [11], it was found that at least in certain circumstances, the motion of the envelopes of these waves were described by a pair of *uncoupled* Nonlinear Schrödinger equations, one for the left moving disturbance and one for the right moving disturbance. Thus, to the degree of approximation usually considered in deriving the amplitude equations, the left and right moving waves behave like "solitons" – they move through each other without interaction.

Recently, building on work in [12], we began to study the analogous question for long wavelength motions on the surface of an incompressible, inviscid fluid. On a formal level, the study of amplitude equations for this problem dates back more than a century to the work of Boussinesq and Korteweg and de Vries and has continued up to the present day. However, surprisingly little was rigorously known about how well these various model equations actually describe fluid motions. To the best of our knowledge, there are only two previous papers in the literature that address this question. In [6], the authors show that for analytic initial conditions that correspond to wave motion in only one direction the solutions of the Euler equation can be approximated by the solution of the Korteweg-de Vries equation. However, a very significant limitation of this work is that it applies only over very short time intervals – time intervals so short that the KdV evolution is essentially trivial. Much closer to our work in spirit is that of Craig [4]. Craig shows that for initial data corresponding to uni-directional motion, the solutions of the water wave problem can be approximated by solutions of the KdV equation for periods of time of the "correct" order of magnitude. (Here "correct" means over times scales of the order one expects on the basis of the formal derivation of the KdV equations.) Our work seeks to go beyond Craig's results in several ways. Most importantly, we wish to consider general long wave length initial data. On physical grounds one expects that such initial data will evolve into two wave packets, one moving to the right and one moving to the left. In light of the discussion of counter-propagating waves in the paragraph above, we will focus in particular on how these two wave packets interact. For a further comparison of our work with that of Craig, we refer the reader to [13].

To describe in more detail our results about the water wave problem, consider the irrotational motion of an incompressible, inviscid fluid in an infinitely long canal of depth one. Choose coordinates $x_1 \in \mathbb{R}$, and x_2 in the bounded direction. Let the free surface be given by $\Gamma(t)$, and let $\Omega(t)$ denote the domain occupied by the fluid at time t. In the parameter regime we study, the fluid's surface $\Gamma(t)$ can always be written as the graph of a function $\eta(x,t)$. The velocity of the fluid $u = (u_1, u_2)$ satisfies Euler's equations:

$$\begin{aligned} \partial_t u + (u \cdot \nabla) u &= -\nabla p + g(0,-1) \ , \\ \nabla \cdot u &= 0 \ . \end{aligned} \tag{1}$$

where p is the pressure and g is the acceleration due to gravity. The boundary conditions appropriate to these circumstances are that $u_2 = 0$ on the bottom of the canal, the pressure is constant on the free surface, and $(1, u_1, u_2)$ is parallel to $(t, \Gamma(t))$.



Initial Conditions: The KdV equation is formally derived in the limit of long-wavelength motion. To insure that the motion we study is of this form, we assume that the initial conditions for the water wave problem are of the form:

$$\eta(x,0) = \epsilon^2 \Phi_1(\epsilon x) \ , \quad u_1(x, \eta(x,0), 0) = \epsilon^2 \Phi_2(\epsilon x) \ . \tag{2}$$

Remark 1.1 It may not be clear that specifying the initial shape of the fluid surface and the horizontal component of the velocity restricted to the initial surface is sufficient to lead to a well posed problem. However, due to the incompressible and irrotational nature of the flow, this turns out to be the case. Indeed, a by-product of our approximation theorem is a long-time existence and uniqueness theorem for solutions of the water-wave problem with long-wavelength initial data.

Let $f(X,T)$ and $g(X,T)$ be solutions of the pair of uncoupled KdV equations

$$\begin{aligned} \partial_T f &= -\frac{1}{6}\partial_X^3 f + \frac{3}{4} f \partial_X f \\ \partial_T g &= \frac{1}{6}\partial_X^3 g - \frac{3}{4} g \partial_X g \end{aligned} \tag{3}$$

with initial conditions $f = \frac{1}{2}(\Phi_2 + \Phi_1)$ and $g = \frac{1}{2}(\Phi_2 - \Phi_1)$.

In informal terms, our result says that the solution of the water wave problem can be approximated by f and g, the solutions of two *uncoupled* KdV equations. To state our result more precisely, we introduce the weighted Sobelev spaces $H^m(n)$, with norm given by $\|u\|_{H^m(n)} = \|(1+|\cdot|^2)^{n/2}u(\cdot)\|_{H^m}$. We also define C_b^n to be the set of functions which are bounded along with their first n derivatives, with $\|u\|_{C_b^n} = \sum_{j=0}^n \|\partial_x^j u\|_{L^\infty}$. We then have:

Theorem 1.2 *Fix $s \geq 5$. For all $C_1, T_0 > 0$, there exist $C_2, \epsilon_0 > 0$ such that for all $\epsilon \in (0, \epsilon_0)$, the following is true. If $\|(\Phi_1, \Phi_2)\|_{H^{s+6}(2)\cap H^{s+10}(2)} \leq C_1$, then there exists a unique solution to the water wave problem with initial conditions (2), and if f and g are the solutions of (3), then one can approximate the solution of the water wave problem as:*

$$\sup_{t\in[0,T_0/\epsilon^3]} \| \begin{pmatrix} \eta \\ u_1 \end{pmatrix} - \epsilon^2 f(\epsilon(\cdot - t), \epsilon^3 t) \begin{pmatrix} 1 \\ -1 \end{pmatrix} - \epsilon^2 g(\epsilon(\cdot + t), \epsilon^3 t) \begin{pmatrix} 1 \\ 1 \end{pmatrix} \|_{C_b^{s-1}} \leq C_2 \epsilon^{5/2}$$

Remark 1.3 In fact, our approximation result contains more information than described here. To save space we have omitted stating the most detailed form of the result. For further details we refer the reader to [13].

The proof of Theorem 1.2 is quite involved, in large part due to the difficulty of establishing the existence and uniqueness of the solutions of the Euler equations. Rather than describe the details of the proof of Theorem 1.2, we will describe in general terms the strategy for proving such approximation theorems, and then work out in detail another case in which the approximation property is similar but the existence theory is much simpler, namely the Fermi-Pasta-Ulam model.

Acknowledgements: The research of CEW was supported in part by the National Science Foundation through grant DMS-9803164.

2 A strategy for proving approximation theorems

In this section we describe the strategy used in [13], and then in the following section we implement that strategy to prove an approximation theorem for solutions of the FPU model.

The general approach we use is due to Kirrmann, Schneider, and Mielke [14]. Suppose that one wishes to study an evolution equation of the form

$$\partial_t u = Lu + B(u,u) \ ,$$

for L a linear operator and B some bilinear term. We will study solutions of small amplitude and suppose that by some formal analysis we have derived an approximate solution which we believe to be close to the true solution. We measure the extent to which it is a good approximation by defining the *residual* to be the amount by which it fails to satisfy the equation, and assuming that the residual is small. More precisely, suppose that our approximate solution is $\epsilon^2\psi$, and that $Res(\epsilon^2\psi) = \epsilon^2(\partial_t\psi - L\psi - \epsilon^2 B(\psi,\psi)) = \mathcal{O}(\epsilon^\gamma)$, for $\gamma > 2$.

We then write the true solution of the problem as $u = \epsilon^2\psi + \epsilon^\beta R$. If we can show that R remains of $\mathcal{O}(1)$ for the times of interest to us and if $\beta > 2$ then we can conclude that $\epsilon^2\psi$ really does provide a good approximation to the true solution. Substituting this form for u into our original equation we see that R satisfies the equation:

$$\partial_t R = LR + 2\epsilon^2 B(\psi, R) + \epsilon^\beta B(R,R) + \epsilon^{-\beta} Res(\epsilon^2\psi) \tag{4}$$

We must address three points in relation to (4).

1. First of all, we will be interested in very long times – here $t \approx \mathcal{O}(1/\epsilon^3)$. We will choose norms such that the linear evolution $\partial_t R = LR$ preserves the norm, but even so, it appears that the linear term $2\epsilon^2 B(\psi, R)$ might cause R to grow uncontrollably over these very long times. In both water wave problem in [13] and the FPU model below, we exploit the form of the non-linearity and the fact that the approximating function ψ has a "long-wavelength" nature to show that in fact, one can bound $\|2\epsilon^2 B(\psi,R)\| \leq C\epsilon^3\|R\|$ in some suitable norm. This allows us to control the linear terms in the equation for the required time interval.

2. Provided the other two conditions are met, we can control the evolution of the nonlinear equation for R using standard "energy methods", if $\beta > 3$. We will choose β to satisfy this inequality.

3. Finally, we must ensure that the inhomogeneous term does not cause R to grow too much. This we will enforce by making sure that the residual satisfies $\|\epsilon^{-\beta} Res(\epsilon^2\psi)\| \leq C\epsilon^3$. To prove this estimate, we take advantage of the fact that we can add to our approximation $\epsilon^2\psi$ any terms of $\mathcal{O}(\epsilon^\mu)$, with $\mu \geq 2$ without effecting the fact that the true solution of the equation is approximated by $\epsilon^2\psi$ to leading order. By choosing these higher order terms appropriately, we will find that we can insure that the residual is small.

Provided that these three points are satisfied, then as explained in [14], (and exploited in a number of other works [10], [11], [12]), an easy application of Gronwall's inequality implies that $R \approx \mathcal{O}(1)$ for $0 < t < \mathcal{O}(1/\epsilon^3)$, and the approximation theorem follows.

3 The Fermi-Pasta-Ulam model

The Fermi-Pasta-Ulam model (FPU) is a system of coupled, nonlinear, Hamiltonian ordinary differential equations:

$$\frac{d^2 q_j}{dt^2} = V'(q_{j+1}(t) - q_j(t)) - V'(q_j(t) - q_{j-1}(t)) \ , \quad j \in \Lambda \ , \tag{5}$$

where $\Lambda \subset \mathbb{Z}$, and $V(q)$ is the potential function for the interparticle forces. The FPU model was first studied numerically by Fermi, Pasta, and Ulam, [8] for a finite set of oscillators in order to see how energy was spread through the various modes of the system by the nonlinear coupling. To their surprise, they found that most trajectories did not "thermalize" as expected, but rather exhibited an almost periodic motion. These observations proved to be extremely fruitful. On one hand, they are related to the Kolmogorov-Arnold-Moser (KAM) theory on the preservation of quasi-periodic motions in nearly integrable Hamiltonian systems. On the other hand, in [16], Kruskal and Zabusky derived the KdV equation as a formal approximation to the FPU model and in studying the KdV equation numerically noticed the fact that solitary waves of that equation seemed to pass through each other without interaction, which they suggested might explain the fact that the FPU model failed to thermalize. This observation lead in turn to the discovery of the complete integrability of the KdV equation. Zabusky and Kruskal had considered initial data corresponding to uni-directional propagation in their derivation of the KdV equation as an approximation to the FPU model. Considering initial data of more general form, Zakharov [17] formally derived the Bousinesq equation as an approximating equation for the FPU lattice, and then discovered that the Bousinesq equation was also completely integrable! A nice description of the history of this model in provided by Bukowski [2]. Kruskal and Zabusky and Zakharov both derived their approximating equations by considering a continuum limit of the FPU model in which the lattice spacing approached zero. The justification of these approximations has been considered in two different circumstances. First, for uni-directional motion, Schwarz [7] studied a system of ODE's related to the FPU model and showed that when solutions of his system of ODE's were appropriately rescaled, they approached solutions of the KdV equation. Closer to our work is that of Bukowski [2] who showed that general motions of the FPU system could be approximated by solutions of the Boussinesq equation for long (but finite) time intervals in the continuum limit. Bukowski's study is complicated by the fact that the Boussinesq equation is ill-posed, so that one must restrict attention to solutions that lie on the center-manifold of the Bousinesq equation. In this section we prove yet a third approximation result for the FPU model. Namely, we show that just as in the case of the water-wave problem, solutions of the FPU model can be approximated as a sum of the solutions of two independent KdV equations, one corresponding to a wave packet moving to the left and one corresponding to a wave packet moving to the right. We present the details of this approximation theorem below because it illustrates many of the issues that arise in the water-wave problem in a situation where there are far fewer technical complications to obscure the main ideas.

We will study an FPU model with infinitely many oscillators, and so we take $\Lambda = \mathbb{Z}$ in (5). In addition, we rewrite (5) in terms of the difference variables $r(j,t) = q_{j+1}(t) - q_j(t)$, so that (5) becomes

$$\partial_t^2 r(j,t) = V'(r(j+1,t)) + V'(r(j-1,t)) - 2V'(r(j,t)) \ , \quad j \in \mathbb{Z} \ . \tag{6}$$

One cannot expect that the KdV equations describe the FPU model under all circumstances, so one must study some asymptotic regime in order to prove this sort of approximation theorem. The one that is usually chosen is the "continuum limit" in which one changes the lattice spacing in (5) from 1 to h, and then lets h tend to zero. A nice formal derivation of the KdV equation as an approximation to the FPU model under this scaling is contained in [1], while a rigorous description of how this leads to the Boussinesq equation is contained in [2]. Note that this rescaling means that a typical structure in the KdV equation (like a soliton, for example,) is spread out over many lattice sites of the FPU model. Another way of achieving this is to keep the lattice spacing in the FPU model fixed at 1, but to rescale the spatial variable in the KdV equation. This is the course we adopt here, in part because it is closer to our approach to the water-wave problem and in part because it also allows us to compare our results with those of Friesecke and Pego [5] who have used a similar rescaling to show that the FPU model possesses travelling wave solutions which remain close *for all time* to a suitably rescaled soliton of the KdV equation. We will refer to this as the "long wavelength" limit of the FPU model, though at least at a formal level it is just another way of looking at the continuum limit.

Let us begin with a formal investigation of the behavior of a "small, long-wavelength" solution. That is, we will assume that the solution of (6) is of the form $r(j,t) = \epsilon^2 R(\epsilon j, t)$, for $R : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$. If we insert this *Ansatz* into (6), and expand V and R with respect to ϵ, we find that to lowest order in ϵ, the function R satisfies $\partial_t^2 R = \epsilon^2 V''(0) \partial_x^2 R$, from which we conclude that at least formally, solutions of (6) split into left and right moving wave packets travelling with speed $\tilde{c} = \epsilon\sqrt{V''(0)}$. This leads to our first hypothesis.

Hypothesis 1 *The potential $V \in C^r$, with $r \geq 5$. Its second derivative satisfies $V''(0) \equiv c^2 > 0$ and its third derivative satisfies $V'''(0) \neq 0$.*

This lowest order calculation leads us to make a more refined "guess" for the form of the solution of (6). We will try to write

$$r(x,t) = \epsilon^2 f(\epsilon(x+ct), \epsilon^3 t) + \epsilon^2 g(\epsilon(x-ct), \epsilon^3 t) + \epsilon^4 \phi(\epsilon x, \epsilon t) \ , \quad x \in \mathbb{Z} \ , \tag{7}$$

with f, g, and ϕ all functions from $\mathbb{R} \times \mathbb{R} \to \mathbb{R}$. Here, f and g represent the left and right moving wave packets, and the dependence on the very slow time scale $\epsilon^3 t$ is included to allow for the effects of the terms that were omitted in the lowest order calculation. The presence of the $\epsilon^4\phi$ term may be more mysterious. These are the "higher order terms" described in point 3 of Section 2. We will choose ϕ in such a way to make sure that the residual of our approximation is small – note that so long as it remains of $\mathcal{O}(1)$ we can choose it however we wish without affecting the lowest order approximation of r by a sum of f and g.

If we now substitute (7) into (6), and expand with respect to ϵ a straightforward computation shows that

$$\begin{aligned}
&\epsilon^4\{c^2\partial_1^2 f(\cdot,\epsilon^3 t) + c^2\partial_1^2 g(\cdot,\epsilon^3 t)\} + \epsilon^6\{2c\partial_1\partial_2 f(\cdot,\epsilon^3 t) - 2c\partial_1\partial_2 g(\cdot,\epsilon^3 t) + \partial_2^2\phi(\epsilon x,\epsilon t)\} + \mathcal{O}(\epsilon^8) \\
&\quad = c^2\{\epsilon^2 f(\cdot+\epsilon,\epsilon^3 t) + \epsilon^2 g(\cdot+\epsilon,\epsilon^3 t) + \epsilon^4\phi(\epsilon x+\epsilon,\epsilon t) + \epsilon^2 f(\cdot-\epsilon,\epsilon^3 t) \\
&\qquad +\epsilon^2 g(\cdot-\epsilon,\epsilon^3 t) + \epsilon^4\phi(\epsilon x-\epsilon,\epsilon t) - 2(\epsilon^2 f(\cdot,\epsilon^3 t) + \epsilon^2 g(\cdot,\epsilon^3 t) + \epsilon^4\phi(\epsilon x,\epsilon t))\} \\
&\quad +\frac{1}{2}V'''(0)\{(\epsilon^2 f(\cdot+\epsilon,\epsilon^3 t) + \epsilon^2 g(\cdot+\epsilon,\epsilon^3 t) + \epsilon^4\phi(\epsilon x+\epsilon,\epsilon t))^2 + (\epsilon^2 f(\cdot-\epsilon,\epsilon^3 t) \\
&\qquad +\epsilon^2 g(\cdot-\epsilon,\epsilon^3 t) + \epsilon^4\phi(\epsilon x-\epsilon,\epsilon t))^2 - 2(\epsilon^2 f(\cdot,\epsilon^3 t) + \epsilon^2 g(\cdot,\epsilon^3 t) + \epsilon^4\phi(\epsilon x,\epsilon t))^2\} \\
&\quad +\frac{1}{3!}V''''(0)\{(\epsilon^2 f(\cdot+\epsilon,\epsilon^3 t) + \epsilon^2 g(\cdot+\epsilon,\epsilon^3 t) + \epsilon^4\phi(\epsilon x+\epsilon,\epsilon t))^3 + (\epsilon^2 f(\cdot-\epsilon,\epsilon^3 t) \\
&\qquad +\epsilon^2 g(\cdot-\epsilon,\epsilon^3 t) + \epsilon^4\phi(\epsilon x-\epsilon,\epsilon t))^3 - 2(\epsilon^2 f(\cdot,\epsilon^3 t) + \epsilon^2 g(\cdot,\epsilon^3 t) + \epsilon^4\phi(\epsilon x,\epsilon t))^3\} \\
&\qquad\qquad +\mathcal{O}(\epsilon^8)
\end{aligned} \tag{8}$$

In the interest of saving space, we have abused notation here – when used as an argument of f, $\cdot = \epsilon(x+ct)$, but when used as an argument of g, $\cdot = \epsilon(x-ct)$. We make a further approximation on the right hand side of (8) by replacing

$$\begin{aligned}
f(\cdot \pm \epsilon, \epsilon^3 t) \quad = \quad & f(\cdot,\epsilon^3 t) \pm \epsilon\partial_1 f(\cdot,\epsilon^3 t) + \frac{1}{2}\epsilon^2\partial_1^2 f(\cdot,\epsilon^3 t) \pm \frac{1}{3!}\epsilon^3\partial_1^3 f(\cdot,\epsilon^3 t) + \frac{1}{4!}\epsilon^4\partial_1^4 f(\cdot,\epsilon^3 t) \\
& \pm\frac{1}{5!}\epsilon^5\partial_1^4 f(\cdot,\epsilon^3 t) + \mathcal{O}(\epsilon^6) \ ,
\end{aligned}$$

with similar expansions for $g(\cdot \pm \epsilon, \epsilon^3 t)$ and $\phi(\epsilon x \pm \epsilon, \epsilon t)$. (This requires a certain amount of smoothness in f and g which we make precise below.) If we insert these expansions into (8), we find that all terms of order ϵ^4 cancel and we are left with:

$$\begin{aligned} &\epsilon^6\{2c\partial_1\partial_2 f - 2c\partial_1\partial_2 g + \partial_2^2\phi\} + \mathcal{O}(\epsilon^8) \\ &\quad = \epsilon^6\{(c^2\frac{1}{12}\partial_1^4 f + \frac{1}{12}\partial_1^4 g + \partial_1^2\phi) + \frac{1}{2}V'''(0)(\partial_1^2(f^2+g^2+2fg))\} + \mathcal{O}(\epsilon^8) \end{aligned} \tag{9}$$

If the "fg" term were not present on the right hand side of (9), then it would be immediately apparent that f and g satisfy a pair of uncoupled KdV equations. While the presence of this term makes it appear that there is an interaction, physically, since f and g are moving in opposite directions, the length of time during which their product will be large will be very small in terms of the slow time scale over which the KdV solutions evolve. Thus one might expect that the effect of the fg term is much smaller than that of either f^2 or g^2. To implement this physical intuition rigorously, we take advantage of our freedom to choose ϕ. Namely, given two solutions f and g of the KdV equation, we will choose ϕ to satisfy

$$\partial_2^2\phi = c^2\partial_1^2\phi + V'''(0)\partial_1^2(fg) \tag{10}$$

That (10) has a well behaved solution is guaranteed by the following lemma.

Lemma 3.1 *Fix $T_0 > 0$ and suppose that $f(\cdot, \epsilon^3 t)$ and $g(\cdot, \epsilon^3 t)$ are elements of $H^{s+1}(2)$, with $s > 2$ for all $0 \le t \le T_0/\epsilon^3$. Then there exists a constant $C_1 > 0$ such that*

$$\sup_{t\in[0,T_0/\epsilon^3]} \|\phi(\cdot,\epsilon t)\|_{H^s} \le C_1 (\sup_{t\in[0,T_0/\epsilon^3]} \|f(\cdot,\epsilon^3 t)\|_{H^{s+1}(2)})(\sup_{t\in[0,T_0/\epsilon^3]} \|g(\cdot,\epsilon^3 t)\|_{H^{s+1}(2)}) ,$$

and $\int(\partial_2\phi(\xi,\epsilon t))d\xi = 0$ for all t.

Proof: See Appendix 1. ■

If we choose ϕ in (7) to have the properties stated in this lemma and repeat the calculations leading to (9), we find that now we have:

$$\begin{aligned} &\epsilon^6\{2c\partial_1\partial_2 f - 2c\partial_1\partial_2 g\} + \mathcal{O}(\epsilon^8) \\ &\quad = \epsilon^6 c^2\{(\frac{1}{12}\partial_1^4 f + \frac{1}{12}\partial_1^4 g) + \frac{1}{2}V'''(0)(\partial_1^2(f^2+g^2))\} + \mathcal{O}(\epsilon^8) \end{aligned} \tag{11}$$

Thus, up to terms of $\mathcal{O}(\epsilon^8)$, we see that the solutions $r(j,t)$ of the FPU model can be formally approximated in the long-wavelength limit by $\epsilon^2 f(\cdot,\epsilon^3 t) + \epsilon^2 g(\cdot,\epsilon^3 t)$ where f and g solve the uncoupled KdV equations:

$$2\partial_2 f = \frac{c}{12}\partial_1^3 f + \frac{V'''(0)}{c} f\partial_1 f \tag{12}$$

$$-2\partial_2 g = \frac{c}{12}\partial_1^3 g + \frac{V'''(0)}{c} g\partial_1 g \tag{13}$$

(14)

In the remainder of this section, we show how to make these formal arguments rigorous.

We first make two further hypotheses concerned with the initial conditions for (6). First note that if we want long-wavelength (and small amplitude) solutions, we expect that $r(j,0) = \epsilon^2\Psi_1(\epsilon j)$ for some function Ψ_1. In addition, our formal computation leads us to expect that the solutions r move with velocities of $\mathcal{O}(\epsilon)$, so we expect that $\partial_t r \approx \mathcal{O}(\epsilon) r$. Our next hypothesis formalizes this intuition.

Hypothesis 2 *Fix constants $C_\Psi, C_0 > 0$ and $s > 2$. We assume that there are functions Ψ_1 and Ψ_2 with $\|\Psi_\ell\|_{H^{s+12}\cap H^7(3)} < C_\Psi$, $(\ell = 1,2)$ such that $\|r(\cdot,0) - \epsilon^2\Psi_1(\epsilon\cdot)\|_{\ell^2(2)} < C_0\epsilon^4$ and $\|\partial_t r(\cdot,0) - \epsilon^3\Psi_2(\epsilon\cdot)\|_{\ell^2(2)} \leq C_0\epsilon^5$.*

Remark 3.2 *Here the Hilbert spaces $\ell^2(m)$ are the discrete analogues of the weighted Sobolev spaces, $H^m(n)$, introduced above. Specifically, $\|r(\cdot)\|^2_{\ell^2(m)} = \sum_{j\in\mathbb{Z}}(1+j^2)^m|r(j)|^2$.*

Finally we note that since $r(j,t) = q_{j+1}(t) - q_j(t)$, it is reasonable to assume that $\sum_{j\in\mathbb{Z}}\partial_t r(j,0) = 0$. Furthermore, if the initial conditions for the FPU model satisfy this conditions, the equations of motion guarantee that it will be satisfied or all time t. Thus, our final hypothesis is:

Hypothesis 3 *Assume that the initial conditions for (6) satisfy $\sum_{j\in\mathbb{Z}}\partial_t r(j,0) = 0$.*

Remark 3.3 Combining hypotheses (2) and (3), with an easy argument involving the Fourier transform of Ψ_2, it is easy to show that there exists a function $\tilde{\Psi}_2^\epsilon \in H^{12}\cap H^7(3)$, and a constant $C_{\tilde{\Psi}}$ with $\|\tilde{\Psi}_2^\epsilon - \Psi_2\|_{H^{12}\cap H^7(3)} \leq C_{\tilde{\Psi}}\epsilon^2$ and $\int\tilde{\Psi}_2^\epsilon(\xi)d\xi = 0$. Furthermore, if $\Xi(\eta) = \int_\infty^\eta \tilde{\Psi}_2^\epsilon(\xi)d\xi$, then $\|\Xi\|_{H^{12}\cap H^7(2)} \leq C_{\tilde{\Psi}}\|\Psi_2\|_{H^{12}\cap H^7(3)}$.

We now state the main theorem of this section.

Theorem 3.4 *Fix $C_\Psi, T_0 > 0$ and suppose that hypotheses 1, 2, and 3 all hold. Then there exists $\epsilon_0, C_1 > 0$ such that for all $\epsilon \in (0,\epsilon_0)$, the solution of (6) satisfies*

$$\sup_{t\in[0,T_0/\epsilon^3]}\|r(\cdot,t) - (\epsilon^2 f(\epsilon(\cdot+ct),\epsilon^3 t) + \epsilon^2 g(\epsilon(\cdot-ct),\epsilon^3 t))\|_{\ell^\infty} \leq C_1\epsilon^{7/2}$$

where f and g are solutions of (12) and (13) respectively, with initial contitions $f(\xi,0) = \frac{1}{2}(\Psi_1(\xi) + \Xi(\xi)/c)$ and $g(\xi,0) = \frac{1}{2}(\Psi_1(\xi) - \Xi(\xi)/c)$.

Remark 3.5 *Note that Hypothesis 2 and Remark 3.3 guarantee that the initial conditions for the KdV equations (12) and (13) are in $H^{12}\cap H^7(2)$. ¿From the existing theory for the KdV equation, one can show that for any $T_0 > 0$, the solutions $f(X,T)$ and $g(X,T)$ of (12) and (13) remain in $H^{12}\cap H^7(2)$ for $0 \leq T \leq T_0$. The details of this argument are included in [13]. This is sufficient smoothness of f and g to make the preceeding formal derivation of the KdV equations rigorous – in fact with more effort one could probably reduce these smoothness requirements further.*

Proof: At various points in the proof it is convenient to work in terms of Fourier transformed variables. Note that if $r:\mathbb{Z}\to\mathbb{R}$, we define $(Fr):[-\pi,\pi]\to\mathbb{C}$ by $(Fr)(p) = \sum_j e^{-ipj}r(j)$. It is often convenient to extend (Fr) periodically to the whole real line. One can invert the Fourier transform by $r(j) = \frac{1}{2\pi}\int_{-\pi}^{\pi}(Fr)(p)e^{ipj}dp$, and hence for a function $\hat{r}\in L^2([-\pi,\pi])$, we define $(F^{-1}\hat{r})(j) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\hat{r}(p)e^{ipj}dp$. Defining $\lambda(p) = c\sqrt{2-2\cos(p)}$, and $(Fs)(p,t) = (\partial_t(Fr)(p,t))/\lambda(p)$, we can rewrite (6) as the first order system of equations

$$\begin{aligned}\partial_t(Fr)(p,t) &= \lambda(p)(Fs)(p,t) &(15)\\ \partial_t(Fs)(p,t) &= -\frac{\lambda(p)}{c^2}(F(V'(r(\cdot,t))))(p,t) &(16)\end{aligned}$$

Remark 3.6 Hypothesis (3) guarantees that $\sum_j\partial_t r(j,t) = 0$, and hence that $(Fs)(p,t)$ is well defined at $p = 0$.

Remark 3.7 Note that $\lambda(p)$ defines a bounded, linear (but non-local) operator on $\ell^2(\mathbb{Z})$ via

$$(\Lambda r)(j) = F^{-1}((\lambda(\cdot)(Fr)(\cdot)))(j)$$

While Λ admits no simple expression, one has $(\Lambda^2 r)(j) = c^2(2r(j) - r(j+1) - r(j-1))$. Note also that if $r(j) = const.$, then $\Lambda r \equiv 0$.

We will also need to use the Fourier transform of functions whose domain is the real numbers rather than the integers. This we define by $(\mathcal{F}f)(p) = \int e^{-ipx} f(x)dx$, with $(\mathcal{F}^{-1}f)(p) = \frac{1}{2\pi}\int e^{ipx} f(x)dx$. Because of our method of approximating the function $r(j,t)$ whose domain is the integers by functions whose domain is the entire real line, it will be useful to have a "truncation operator", $\mathcal{T} : L^2(\mathbb{R}) \to \ell^2(\mathbb{Z})$, which we define by $(\mathcal{T}h) = F^{-1}(\chi_{[-\pi,\pi]}\mathcal{F}h)$, where $\chi_{[-\pi,\pi]}$ is the characteristic function of the interval $[-\pi,\pi]$. Note that $\mathcal{T}$ has norm one. If $h \in C^0(\mathbb{R})$, the other natural way to represent it as a function on the integers is simply by restricting its domain. That these two approaches yield similar results, at least for "long wave-length" functions is guaranteed by the following lemma.

Lemma 3.8 *Fix $m \geq 1$. There exists a constant $C_m > 0$ such that if $h(x) = \mathcal{H}(\epsilon x)$, with $\mathcal{H} \in H^m$, then*

$$\|h|_{\mathbb{Z}} - \mathcal{T}h\|_{\ell^2} \leq C_m \epsilon^{m-\frac{1}{2}} \|\mathcal{H}\|_{H^m}$$

Proof: The proof is an easy exercise with Fourier transforms which we complete in Appendix 2 ■

We will also need to estimate the action of Λ on "long-wavelength" functions.

Lemma 3.9 *Suppose in the following that $h(x) = \mathcal{H}(\epsilon x)$, for $\mathcal{H} \in H^m$, $m \geq 2$. Then one has the following estimates.*

$$\|h|_{\mathbb{Z}}\|_{\ell^2} \leq \epsilon^{-1/2}\|\mathcal{H}\|_{H^m} \tag{17}$$

$$\|\Lambda h|_{\mathbb{Z}}\|_{\ell^2} \leq \epsilon^{1/2}\|\mathcal{H}\|_{H^m} \tag{18}$$

Proof: Once again, this in an exercise in Fourier tranforms which we defer to the appendix. ■

Following the strategy outlined in Section 2, we write

$$r(j,t) = \epsilon^2\psi(j,t) + \epsilon^{7/2}R(j,t)\ , \quad j \in \mathbb{Z} \tag{19}$$

where $\psi(x,t) = f(\epsilon(x+ct), \epsilon^3 t) + g(\epsilon(x-ct), \epsilon^3 t) + \epsilon^2\phi(\epsilon x, \epsilon t)$ is actually defined for all $x \in \mathbb{R}$ – not just for $j \in \mathbb{Z}$. In like fashion, we write s as

$$\begin{aligned} s(j,t) &= \epsilon^2 F^{-1}\left(\frac{1}{\lambda(p)}\chi_{[-\pi,\pi]}(p)(\mathcal{F}\partial_t(\psi(\cdot,t)))\right)(j,t) + \epsilon^{7/2}S(j,t) \\ &= \epsilon^2(\Lambda^{-1}\mathcal{T}\partial_t\psi)(j,t) + \epsilon^{7/2}S(j,t)\ . \end{aligned} \tag{20}$$

The reason for this somewhat complicated definition is as follows. The natural approximation for (Fs) would be $\epsilon^2(F(\partial_t\psi))(p)/\lambda(p)$. However, we cannot be sure that that $F(\partial_t\psi)(0) = 0$, and hence we choose to approximate $F(\partial_t\psi)(p)$ by $(\mathcal{F}\partial_t(\psi(p,t)))$. We see from its definition that this quantity vanishes at $p = 0$ and hence $\frac{1}{\lambda(p)}(\mathcal{F}dt\psi)(p,t)$ is well defined.

The proof of Theorem 3.4 is completed by showing that if R and S are of $\mathcal{O}(1)$ at time $t=0$, they remain so for all $t \leq T_0/\epsilon^3$. From (15), we see that R and S satisfy the system of equations

$$\partial_t R(j,t) = (\Lambda S)(j,t) + \epsilon^{2-7/2}(\mathcal{T}(\partial_t\psi))(j,t) - \partial_t\psi(j,t)) \quad (21)$$

$$\partial_t S(j,t) = \epsilon^{-7/2}(-\frac{1}{c^2}\Lambda V'(\epsilon^2\psi + \epsilon^{7/2}R))(j,t) - \epsilon^{2-7/2}(\partial_t\Lambda^{-1}\mathcal{T}(\partial_t\psi))(j,t) \quad (22)$$

We will bound the solutions of (21) in the norm

$$\|(R,S)\|^2 = \sum_{j\in\mathbb{Z}}(R(j)^2 + S(j)^2) + \frac{\epsilon^2}{2c^2}V'''(0)\sum_{j\in\mathbb{Z}}\psi(j,t)R(j)^2 \ .$$

Remark 3.10 If f and g are in $H^4(2)$, it is easy to show that for all times $0 \leq t \leq T_0/\epsilon^3$, there exists c_1 and c_2 such that for ϵ sufficiently small,

$$c_1(\|R\|_{\ell^2} + |S\|_{\ell^2}) \leq \|(R,S)\| \leq c_2(\|R\|_{\ell^2} + |S\|_{\ell^2})$$

so the $\|(\cdot,\cdot)\|$ norm is equivalent to the ℓ^2 norm on each component. Note too that $\|R\|_{\ell^\infty} \leq \|R\|_{\ell^2}$.

If we define $Res_1(\epsilon^2\psi) = \epsilon^2(\mathcal{T}\partial_t(\psi))(j,t) - \partial_t\psi(j,t))$, and $Res_2(\epsilon^2\psi) = -\frac{1}{c^2}\Lambda V'(\epsilon^2\psi) - \epsilon^2(\partial_t\Lambda^{-1}\mathcal{T}\partial_t(\psi))$, then again following the strategy outlined in Section 2, we can expand $V'(\epsilon^2\psi + \epsilon^{7/2}R)$ and rewrite (21) as

$$\partial_t R = \Lambda S + \epsilon^{-7/2}Res_1(\epsilon^2\psi) \quad (23)$$

$$\partial_t S = -\Lambda R - \frac{\epsilon^2}{2c^2}V'''(0)\Lambda\psi R + G(R,\psi) + \epsilon^{-7/2}Res_2(\epsilon^2\psi) \quad (24)$$

where given any constant C_R, there exists a constant C_G, such that $\|G(R,\psi)\|_{\ell^2} \leq C_G(\epsilon^3\|R\|_{\ell^2} + \epsilon^{7/2}\|R\|_{\ell^2}^2)$, for all $\|R\|_{\ell^2} \leq C_R$.

¿From Lemma 3.8, we see that if f and g are in H^{11}, then $\|Res_1(\epsilon^2\psi)\|_{\ell^2} \leq C\epsilon^{13/2}$. (To see why we apparently lose more derivatives than required by Lemma 3.8, note that $\partial_2 f = \frac{c}{12} + f(\partial_1 f) \in H^8$, and similarly for g.)

Turning to $Res_2(\epsilon^2\psi)$, Note that $V'(\epsilon^2\psi) = V'(0) + c^2\epsilon\psi + \frac{1}{2}V'''(0)(\epsilon\psi)^2 + R_3(\epsilon\psi))$, where $R_3 \approx (\epsilon\psi)^3$ By Lemma 3.9, $\|\Lambda R_3\|_{\ell^2} \leq C\epsilon^{13/2}$, and by Remark 3.7, $\Lambda V'(0) = 0$. Next note that $(\epsilon^2\psi)^2 = \epsilon^4(f^2 + g^2 + 2fg) + 2\epsilon^6(f+g)\phi + \epsilon^8\phi^2$. Again, using Lemma 3.9, we note that $\|\Lambda(2\epsilon^6(f+g)\phi + \epsilon^8\phi^2)\|_{\ell^2} \leq C\epsilon^{13/2}$.

The only remaining terms in Res_2 are $-\epsilon^2\Lambda(f+g+\epsilon^2\phi) - \frac{\epsilon^4 V'''(0)}{2c^2}\Lambda(f^2+g^2+2fg) - \epsilon^2\Lambda^{-1}\mathcal{T}(\partial_t^2(f+g+\epsilon^2\phi))$. Recalling the formal calculations earlier in this section which lead to the definitions of f, g and ϕ, we expect that these terms are small. That the formal calculations are correct is guaranteed by the following lemma.

Lemma 3.11 *If f and g are elements of $H^{11} \cap H^7(2)$, and satisfy (12) and (13), and ϕ is as in Lemma 3.1, then*

$$\|-\epsilon^2\Lambda(f+g+\epsilon^2\phi) - \frac{\epsilon^4 V'''(0)}{2c^2}\Lambda(f^2+g^2+2fg) - \epsilon^2\Lambda^{-1}\mathcal{T}(\partial_t^2(f+g+\epsilon^2\phi))\|_{\ell^2} \leq C\epsilon^{13/2} \ . \quad (25)$$

We will prove this lemma in Appendix 3, and continue here with the proof of Theorem 3.4. The estimates in the preceeding paragraphs imply that the ℓ^2 norm of both Res_1 and Res_2 can be bounded by $C\epsilon^{13/2}$. With these estimates in hand, we differentiate the norm $\|(R,S)\|^2$ along a solution and find that

$$\begin{aligned}\frac{d}{dt}\|(R,S)\|^2 = 2\sum_j R(j,t)(\epsilon^{-7/2}Res_1(\epsilon^2\psi)(j,t)) + 2\sum_j S(j,t)[(-\frac{\epsilon^2}{2c^2}\Lambda V'''(0)\psi R)(j,t) \\ +G(R,\psi)(j,t) + \epsilon^{-7/2}Res_2(\epsilon^2\psi)(j,t)] \\ +\frac{\epsilon^2}{2c^2}V'''(0)\sum_{j\in\mathbb{Z}}(\partial_t\psi(j,t))R(j,t)^2 + \frac{\epsilon^2}{c^2}V'''(0)\sum\psi(j,t)R(j,t)[(\Lambda S)(j,t) \\ +\epsilon^{-7/2}Res_1(\epsilon^2\psi)(j,t)]\ .\end{aligned}$$

Note that the terms involving the residues, Res_1 and Res_2 can be bounded by $C\epsilon^3(\|(R,S)\|)$. Similarly, if we differentiate ψ with respect to time, we pick up a factor of ϵ, so that we can bound $|\frac{\epsilon^2}{2}V'''(0)\sum_{j\in\mathbb{Z}}(\partial_t\psi(j,t))S(j,t)^2| \leq C\epsilon^3\|(R,S)\|^2$. Given the estimates on $G(R,\psi)$ above, for any $C_R > 0$, we can find $C_G > 0$ such that we can bound the term involving G by $C_G(\epsilon^3\|R\|_{\ell^2} + \epsilon^{7/2}\|R\|^2_{\ell^2})\|(R,S)\|$, for all $\|R\|_{\ell^2} \leq C_R$. Finally the two remaining terms of $\mathcal{O}(\epsilon^2)$ cancel.

Combining all the estimates in the previous paragraph, we see that for any $C_R > 0$, we can bound

$$\begin{aligned}\frac{d}{dt}\|(R,S)\|^2 &\leq C_1\epsilon^3\|(R,S)\|^2 + C_2\epsilon^{7/2}\|(R,S)\|^2 + C_3\epsilon^3\|(R,S)\| \\ &\leq C_1\epsilon^3\|(R,S)\|^2 + C_2\epsilon^{7/2}\|(R,S)\|^2 + 2(C_3\epsilon^3 + \epsilon^3\|(R,S)\|^2)\end{aligned} \tag{26}$$

for all $\|(R,S)\| \leq C_R$. Applying Gronwall's inequality to this estimate, we see that provided the initial conditions $(R(j,0),S(j,0))$ are sufficiently small, the solutions $(R(\cdot,t),S(\cdot,t))$ of (21) are bounded by $2C_3\exp(C_1T_0)$, for all $0 \leq t \leq T_0/\epsilon^3$. However, both components of the initial conditions have small ℓ^2 norm by Hypothesis 2 and equations (19) and (20). Since the ℓ^∞ norm is bounded by a constant times the $\|(\cdot,\cdot)\|$ norm, this completes the proof of Theorem 3.4. ■

Acknowledgements: The work of CEW was supported in part by the National Science Foundation through grant DMS-9803164.

Appendix 1 Proof of Lemma 3.1

Recall that $\phi = \phi(\xi,\tau)$ satisfies (10) which we rewrite as the first order system

$$\begin{aligned}\partial_2\phi &= \psi \\ \partial_2\psi &= \partial_1^2\phi + V'''(0)\partial_1^2(fg)\end{aligned} \tag{27}$$

Taking Fourier transforms, and writing out the explicit form of the semi-group of the homogeneous part of the equation, we obtain the (Fourier transform of a) particular solution of (27)

$$\begin{pmatrix}(\mathcal{F}\phi)(p,\tau) \\ (\mathcal{F}\psi)(p,\tau)\end{pmatrix} = V'''(0)\int_0^\tau \begin{pmatrix}-p\sin(p(\tau-s))(\mathcal{F}(fg))(p,s) \\ -p^2\cos(p(\tau-s))(\mathcal{F}(fg))(p,s)\end{pmatrix} ds$$

The first thing we see from this formula is that $(\mathcal{F}\psi)(0,\tau) = 0$ for all τ, from which we conclude that $\int(\partial_2\phi)(\xi,\tau)d\xi = 0$ for all τ.

Next note that since

$$(\mathcal{F}\phi)(p,\tau) = V'''(0)\int_0^\tau -p\sin(p(\tau-s))(\mathcal{F}(fg))(p,s)ds$$

we have

$$\begin{aligned}
\|\phi(\cdot,\tau)\|_{H^s} &\leq \|(\mathcal{F}\phi)(\cdot,\tau)\|_{H^0(m)} \\
&\leq C_1\int_0^\tau \|(\mathcal{F}(fg))(\cdot,s)\|_{H^0(m+1)}d\tau \leq C_1\int_0^\tau \|(fg)(\cdot,s)\|_{H^{m+1}}d\tau \\
&\leq C_2\int_0^\tau \sup_x\left(\frac{1}{(1+|x+cs|)^2}\frac{1}{(1+|x-cs|)^2}\right)ds \\
&\qquad \times \sup_{s\in[0,\tau]} \|f(\cdot,\epsilon^2 s)\|_{H^{m+1}(2)}\|g(\cdot,\epsilon^2 s)\|_{H^{m+1}(2)} \\
&\leq C_3\|f(\cdot,\epsilon^2 s)\|_{H^{m+1}(2)}\|g(\cdot,\epsilon^2 s)\|_{H^{m+1}(2)} \qquad (28)
\end{aligned}$$

The estimate in Lemma 3.1 follows immediately from (28).

Appendix 2 Proof of Lemmas 3.8 and 3.9

Proof: (of Lemma 3.8) By Parseval's identity, we have

$$\|h|_{\mathbb{Z}} - \mathcal{T}h\|_{\ell^2}^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} |(Fh|_{\mathbb{Z}})(p) - (F(\mathcal{T}h))(p)|^2 dp$$

By explicit calculation one finds that $(Fh|_{\mathbb{Z}})(p) = \frac{1}{\epsilon}\sum_{n\in\mathbb{Z}}(\mathcal{F}\mathcal{H})(\frac{p+2\pi n}{\epsilon})$, while $(F(\mathcal{T}h))(p) = (\mathcal{F}h)(p) = \frac{1}{\epsilon}(\mathcal{F}\mathcal{H})(\frac{p}{\epsilon})$ by definition. Thus,

$$\begin{aligned}
\|h|_{\mathbb{Z}} - \mathcal{T}h\|_{\ell^2}^2 &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \|\frac{1}{\epsilon}\sum_{n\neq 0}(\mathcal{F}\mathcal{H})(\frac{p+2\pi n}{\epsilon})|^2 dp \\
&\leq \frac{1}{2\pi}\sum_{n_1,n_2\neq 0}\int_{-\pi}^{\pi}\frac{1}{\epsilon^2}\frac{(1+|\frac{p+2\pi n_1}{\epsilon}|^2)^{m/2}}{(1+|\frac{p+2\pi n_1}{\epsilon}|^2)^{m/2}}\frac{(1+|\frac{p+2\pi n_2}{\epsilon}|^2)^{m/2}}{(1+|\frac{p+2\pi n_2}{\epsilon}|^2)^{m/2}} \\
&\qquad \times\|(\mathcal{F}\mathcal{H})(\frac{p+2\pi n_1}{\epsilon})\|\|(\mathcal{F}\mathcal{H})(\frac{p+2\pi n_1}{\epsilon})|dp \\
&\leq \frac{C}{\epsilon}\sum_{n_1,n_2\neq 0}\frac{1}{(1+|\frac{\pi n_1}{\epsilon}|^2)^{m/2}}\frac{1}{(1+|\frac{\pi n_2}{\epsilon}|^2)^{m/2}}\|\mathcal{H}\|_{H^m}^2 \leq C\epsilon^{2m-1}\|\mathcal{H}\|_{H^m}^2
\end{aligned}$$

■

Proof: (of Lemma 3.9)

$$\begin{aligned}
\|h|_{\mathbb{Z}}\|_{\ell^2}^2 &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\|(Fh|_{\mathbb{Z}})(p)|^2 dp \\
&\leq \frac{1}{2\pi}\sum_{n_1,n_1}\int_{-\pi}^{\pi}\|(\mathcal{F}\mathcal{H})(\frac{p+2\pi n_1}{\epsilon})\|\|(\mathcal{F}\mathcal{H})(\frac{p+2\pi n_1}{\epsilon})|dp \\
&\leq \frac{C}{\epsilon}\sum_{n_1,n_2}\frac{1}{(1+|\frac{\pi n_1}{\epsilon}|^2)^{m/2}}\frac{1}{(1+|\frac{\pi n_2}{\epsilon}|^2)^{m/2}}\|\mathcal{H}\|_{H^m}^2 \leq C\epsilon^{-1}\|\mathcal{H}\|_{H^m}^2
\end{aligned}$$

where we used the observation of the previous paragraph to rewrite $(Fh|_{\mathbb{Z}})$ in terms of $(\mathcal{FH})$. This completes the proof of (17) The proof of (18) goes along similar lines. One has

$$\begin{aligned}
\|\Lambda h|_{\mathbb{Z}}\|_{\ell^2}^2 &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\|\lambda(p)(Fh|_{\mathbb{Z}})(p)|^2 dp \\
&\le \frac{c^2}{2\pi}\sum_{n_1,n_1}\int_{-\pi}^{\pi}(2\cos(p)-2)\|(\mathcal{FH})(\frac{p+2\pi n_1}{\epsilon})\|\|(\mathcal{FH})(\frac{p+2\pi n_1}{\epsilon})|dp \\
&\le \frac{c^2}{2\pi}\sum_{n_1,n_1}\frac{1}{(1+|\frac{\pi n_1}{\epsilon}|^2)^{(m-1)/2}}\frac{1}{(1+|\frac{\pi n_2}{\epsilon}|^2)^{(m-1)/2}} \\
&\qquad \times(\int_{-\pi/\epsilon}^{\pi/\epsilon}(2\cos(\epsilon q)-2)(1+q^2)^{m-1}|\mathcal{FH}(q)|^2 dq) \\
&\le \frac{C}{2\pi\epsilon}\int_{-\infty}^{\infty}\epsilon^2 q^2(1+q^2)^{m-1}|\mathcal{FH}(q)|^2 dq \le C\epsilon\|\mathcal{H}\|_{H^m}^2
\end{aligned}$$

■

Appendix 3 Proof of Lemma 3.11

Consider the terms in (25) depending on f alone – *i.e.* $-\epsilon^2\Lambda f - \frac{\epsilon^4 V'''(0)}{2c^2}\Lambda f^2 - \epsilon^2\Lambda^{-1}(\mathcal{T}\partial_t^2 f)$ Recalling that $f = f(\epsilon(x+ct),\epsilon^3 t)$, and that f is a solution of (12), we see that $\partial_t^2 f = c^2\epsilon^2\partial_1^2 f + 2c\epsilon^4\partial_1\partial_2^2 f + \epsilon^6\partial_2 f$, where to save space, we have supressed the arguments of f. From (13), we see that $4\partial_2^2 f = \partial_1^6 f + \partial_1^4(f^2) + 2\partial_1(f(\partial_1^3 f + \partial_1 f))$. Taking Fourier transforms, we see that if $f \in H^{11}$, we can immediately bound $\|\epsilon^8\Lambda^{-1}\partial_2^2 f\|_{\ell^2} \le C\epsilon^{13/2}$, so we can ignore this term. (Note that the powers of p in the Fourier transform of this expression cancel the possible divergence coming from the $1/\lambda(p)$ near $p=0$.) To bound the remaining terms, note that taking Fourier transforms we have

$$\begin{aligned}
&F(-\epsilon^2\Lambda f - \frac{\epsilon^4 V'''(0)}{2c^2}\Lambda f^2 - \epsilon^4\Lambda^{-1}\mathcal{T}(c^2\partial_1^2 f + 2c\epsilon^2\partial_1\partial_2 f)) = \\
&= \lambda(p)(-\epsilon^2(Ff) - \frac{\epsilon^4 V'''(0)}{2c^2}(F(f^2))) - \epsilon^4\frac{1}{\lambda(p)}\chi_{[-\pi,\pi]}(-\frac{c^2}{\epsilon}(\frac{p}{\epsilon})^2(\mathcal{F}f) \\
&\quad +2c\epsilon^2(\frac{c}{12\epsilon}(\frac{p}{\epsilon})^4(\mathcal{F}f) - (\frac{p}{\epsilon})^2\frac{V'''(0)}{2c\epsilon}^2(\mathcal{F}(f^2))))
\end{aligned} \tag{29}$$

Note that the factors of $(\frac{p}{\epsilon})^n$ occur because of the fact that f is a function of ϵx. Also, we have abused notation again here in that since f is really a function of $f(\epsilon(x+ct),\epsilon^3 t)$, we get not $\mathcal{F}f$, but $e^{-ipct/\epsilon}\mathcal{F}f$. However these factors of absolute value one have no effect on the final estimate so we have omitted them to avoid cluttering the notation unncessarily. Because f is of "long wavelength" form, we can use the results of Appendix 2 to express Ff in terms of $\mathcal{F}f$. ¿From this we see that we can rewrite

$$-\lambda(p)(\epsilon^2(Ff) + \frac{\epsilon^4 V'''(0)}{2c^2}(F(f^2))) = -\lambda(p)(\epsilon^2\frac{1}{\epsilon}(\mathcal{F}f)(\frac{p}{\epsilon}) + \frac{\epsilon^4 V'''(0)}{2c^2}(\mathcal{F}(f^2))(\frac{p}{\epsilon})) + \mathcal{E}_{res}$$

where by the same methods used in the proof of Lemma 3.9 we have the estimate $\|F^{-1}\mathcal{E}_{res}\|_{\ell^2} \le C\epsilon^{13/2}$, since $f \in H^{11}$ (In fact, for this estimate, it would suffice for f to be in H^7.). Thus, ignoring $\mathcal{E}_{res}$, we find we must bound

$$\| - \lambda(p)(\epsilon^2 \frac{1}{\epsilon}(\mathcal{F}f)(\frac{\cdot}{\epsilon}) + \frac{\epsilon^4 V'''(0)}{2c^2}(\mathcal{F}(f^2))(\frac{\cdot}{\epsilon})) \tag{30}$$

$$-\epsilon^4 \frac{1}{\lambda(p)}(-\frac{c^2}{\epsilon}(\frac{\cdot}{\epsilon})^2(\mathcal{F}f)(\frac{\cdot}{\epsilon}) + c\epsilon^2[\frac{c}{12\epsilon}(\frac{\cdot}{\epsilon})^4(\mathcal{F}f)(\frac{\cdot}{\epsilon}) - (\frac{\cdot}{\epsilon})^2 \frac{V'''(0)}{2c\epsilon}^2 (\mathcal{F}(f^2))(\frac{\cdot}{\epsilon})])\|^2_{L^2[-\pi,\pi]}$$

Writing out the integral corresponding to this norm, and changing variables in the integral from p to $q = p/\epsilon$, we find that we must estimate

$$\epsilon \int_{-\pi/\epsilon}^{\pi/\epsilon} |\{(\lambda(\epsilon q))^2[-\epsilon(\mathcal{F}f)(q) - \frac{\epsilon^3}{2c^2}V'''(0)(\mathcal{F}(f^2))(q)] \tag{31}$$

$$+\epsilon^3 c^2 q^2(\mathcal{F}f)(q) - \frac{\epsilon^5 c^2}{12}q^4(\mathcal{F}f)(q) + \frac{\epsilon^5 V'''(0)}{2}q^2(\mathcal{F}f)(q)\}/\lambda(\epsilon q)|^2 dq$$

Recall that $(\lambda(\epsilon q))^2 = 2c^2(1 - \cos(\epsilon q)) = c^2((\epsilon q)^2 - \frac{c^2}{12}(\epsilon q)^4) + \mathcal{O}(\epsilon^6)$. But this means that

$$(\lambda(\epsilon q))^2[-\epsilon(\mathcal{F}f)(q)] + (\epsilon^3 c^2 q^2 - \frac{\epsilon^5 c^2}{12}q^4)(\mathcal{F}f)(q) = \mathcal{O}(\epsilon^7 q^6)(\mathcal{F}f)(q)$$

while

$$-(\lambda(\epsilon q))^2 \frac{\epsilon^3 V'''(0)}{2c^2}(\mathcal{F}f^2)(q) + \frac{\epsilon^5 V'''(0)}{2}q^2(\mathcal{F}f^2)(q) = \mathcal{O}(\epsilon^7 q^3)(\mathcal{F}f^2)(q) \ .$$

Since $|\frac{\epsilon q}{\lambda(q)}| \leq 10/c$ for $q \in [-\pi/\epsilon, \pi/\epsilon]$, we can bound (31) by $C\epsilon^{13}(\|(1+q^2)^4\mathcal{F}f\|^2_{L^2} + \|(1+q^2)^2(\mathcal{F}f^2)\|^2_{L^2} \leq C\epsilon^{13}(\|f\|^2_{H^8} + \|f\|^4_{H^8})$. Thus, so long as $f \in H^8 \subset H^{11}$, we have

$$\| - \epsilon^2 \Lambda f - \frac{\epsilon^4 V'''(0)}{2c^2}\Lambda f^2 - \epsilon^2 \Lambda^{-1}(\mathcal{T}\partial_t^2 f)\|_{\ell^2} \leq C\epsilon^{13/2} \ .$$

The terms involving g and ϕ are handled in a very similar fashion, which completes the proof of Lemma 3.11.

References

[1] M. Ablowitz and H. Segur. *Solitons and the inverse scattering transform.* SIAM Studies in Applied Mathematics. SIAM, Philadelphia, 1981.

[2] John Bukowksi. *The Boussinesq limit of the Fermi-Pasta-Ulam Equation.* PhD thesis, Brown University, 1997.

[3] Pierre Collet and Jean-Pierre Eckmann. The amplitude equation and the Swift-Hohenberg equation. *Comm. Math. Phys.*, 132:139–153, 1990.

[4] Walter Craig. An existence theory for water-waves and the Boussinesq and Korteweg-de Vries scaling limits. *Comm. Part. Diff. Equations*, 10:787–1003, 1985.

[5] G. Friesecke and R. Pego. Solitary waves on FPU lattices I: Qualitative properties, renormalization and continuum limit. Preprint: to appear in Nonlinearity, 1999.

[6] T. Kano and T. Nishida. A mathematical justification for Korteweg–de Vries equation and Boussinesq equation of water surface waves. *Osaka J. Math.*, 23:389–413, 1986.

[7] Jr. Martin Schwarz. Korteweg-de Vries and nonlinear equations related to the Toda lattice. *Advances in Mathematics*, 44:132–154, 1982.

[8] E. Fermi J. Pasta and S. Ulam. Studies of nonlinear problems, I. Technical Report Los Alamos Rep, LA1940, Los Alamos, 1955. reproduced in *Nonlinear Wave Motion*, A.C. Newell, ed. AMS, Providence, RI, 1974.

[9] Robert Pierce and C. Eugene Wayne. On the validity of mean-field amplitude equations for counterpropagating wavetrains. *Nonlinearity*, 8:769–779, 1995.

[10] Guido Schneider. Error estimates for the Ginzburg-Landau approximation. *Zeit. Angew. Math. Phys. (ZAMP)*, 45:433–457, 1994.

[11] Guido Schneider. Justification of mean-field coupled modulation equations. *Proc. Roy. Soc. Edinburgh Sect. A*, 127:639–650, 1997.

[12] Guido Schneider. The long wave limit for a Boussinesq equation. *SIAM Journal of Applied Mathematics*, 58:1237–1245, 1998.

[13] Guido Schneider and C. Eugene Wayne. The long wave limit for the water wave problem, I: The case of zero surface tension. preprint, 1999.

[14] P. Kirrmann G. Schneider and A. Mielke. The validity of modulation equations for extended systems with cubic nonlinearities. *Proc. R. Soc. Edin.*, 122A:85–91, 1992.

[15] A. van Harten. On the validity of Ginzburg–Landau's equation. *J. Nonlinear Science*, 1:397–422, 1991.

[16] N.J. Zabusky and M.D. Kruskal. Interactions of solitons in a collisionless plasma and the recurrence of initial states. *Phys. Rev. Lett.*, 15:240–243, 1965.

[17] V. E. Zakharov. On stochastization of one-dimensional chains of nonlinear oscillators. *Sov. Phys. JETP*, 38:108–110, 1974.