Commun. Math. Phys. 164, 305–349 (1994)

Communications in
Mathematical
Physics

Asymptotic Stability of Solitary Waves

Robert L. Pego[1], **Michael I. Weinstein**[2]

[1] Department of Mathematics & Institute for Physical Science and Technology, University of Maryland, College Park, MD 20742, USA. Supported by NSF Grants DMS 9196155 and 9201869.
[2] Department of Mathematics, University of Michigan, Ann Arbor, MI 48109, USA. Supported by NSF Grant DMS 9201717.

Received: 20 August 1993

Abstract: We show that the family of solitary waves (1-solitons) of the Korteweg–de Vries equation

$$\partial_t u + u\partial_x u + \partial_x^3 u = 0 ,$$

is asymptotically stable. Our methods also apply for the solitary waves of a class of generalized Korteweg–de Vries equations,

$$\partial_t u + \partial_x f(u) + \partial_x^3 u = 0 .$$

In particular, we study the case where $f(u) = u^{p+1}/(p+1)$, $p = 1, 2, 3$ (and $3 < p < 4$, for $u > 0$, with $f \in C^4$). The same asymptotic stability result for KdV is also proved for the case $p = 2$ (the modified Korteweg–de Vries equation). We also prove asymptotic stability for the family of solitary waves for all but a finite number of values of p between 3 and 4. (The solitary waves are known to undergo a transition from stability to instability as the parameter p increases beyond the critical value $p = 4$.) The solution is decomposed into a modulating solitary wave, with time-varying speed $c(t)$ and phase $\gamma(t)$ (*bound state part*), and an infinite dimensional perturbation (*radiating part*). The perturbation is shown to decay exponentially in time, in a local sense relative to a frame moving with the solitary wave. As $p \to 4^-$, the local decay or radiation rate decreases due to the presence of a *resonance pole* associated with the linearized evolution equation for solitary wave perturbations.

1. Introduction

Solitary waves are a class of finite energy, spatially localized solutions of nonlinear dispersive partial differential equations of Hamiltonian type. In many such systems, computer simulations and certain analytical results suggest that, in general, solutions eventually resolve themselves into an approximate superposition of weakly interacting solitary waves and decaying dispersive waves. Thus it has been suggested (see [L, GGKM1] that solitary waves play the role of elements in a *nonlinear basis*, with respect to which it is natural to view the solution in the limit of large time. A natural step toward understanding this sort of asymptotic decomposition is to consider the stability of solitary waves. This is the study of the behavior of solutions with initial conditions in a neighborhood of the solitary wave.

In this paper we establish a result concerning the asymptotic stability of solitary wave solutions of the Korteweg–de Vries (KdV) equation

$$\partial_t u + u\partial_x u + \partial_x^3 u = 0 . \tag{1.1}$$

The methods we use also apply, for example, to the solitary waves of a generalized KdV equation (gKdV)

$$\partial_t u + \partial_x f(u) + \partial_x^3 u = 0 . \tag{1.2}$$

In particular, we study the case where $f(u) = u^{p+1}/(p+1)$, for $p = 1, 2, 3$, (and $3 < p < 4$ for $u > 0$, with $f \in C^4$). The same asymptotic stability result which is proved for KdV is shown to hold for the case, $p = 2$, the modified Korteweg–de Vries equation (mKdV). The solitary waves of (1.2) are known to undergo a transition from stability to instability as the parameter p increases beyond the critical value $p = 4$, cf. [LS, W1, W3, BSS, PW2]. Some of the results of the present paper were announced in [PW1].

The KdV and gKdV equations have a two-parameter family of solitary wave solutions of the form $u(x, t) = u_c(x - ct + \gamma)$, for all $c > 0$, $\gamma \in \mathbb{R}$. The solitary wave profile $u_c(y)$ is the unique symmetric solution of the equation

$$-\partial_y^2 u_c + cu_c - f(u_c) = 0 , \tag{1.3}$$

having $u_c(y) \to 0$ as $|y| \to \infty$. Explicitly, for our particular nonlinearity, we have

$$u_c(y) = \alpha \operatorname{sech}^{2/p} \beta y, \quad \text{where } \alpha = \left(\frac{1}{2}c(p+1)(p+2)\right)^{1/p}, \quad \beta = \frac{1}{2}p\sqrt{c} . \tag{1.4}$$

Because a small perturbation of a solitary wave can yield a solitary wave with a permanent phase shift, or one with a different speed, it is appropriate to study the *orbital stability* of solitary waves. An extensive mathematical literature on the subject of orbital stability of solitary waves developed following the work of Benjamin [Be] (see also Bona [Bo]) for the KdV equation. The results of Laedke and Spatschek [LS], Weinstein [W1], [W3] and Bona, Souganidis and Strauss [BSS] (see also [BSo]) assert that for integer p with $1 \leq p < 4$, a solution which is initially close to a solitary wave $u_c(x - ct)$ in the Sobolev space $H^1(\mathbb{R})$, will forever remain close to the set of translates $u_c(x - ct + \gamma)$ of the wave. (This is the orbit of the wave under the group of time translations.) Somewhat more precisely, for sufficiently small $\delta > 0$, one has

$$\inf_\gamma \| u(\cdot, t) - u_c(\cdot + \gamma) \|_{H^1} < \delta \tag{1.5}$$

for all $t > 0$, if the same quantity is small at the initial time $t = 0$. This notion of stability establishes that the *shape* of the wave is stable, but does not fully resolve the question of what the asymptotic behavior of the system is. *Á priori*, it is possible that the solution wanders in a neighborhood of the group orbit of the solitary wave without settling down to some well-defined asymptotic state.

Our goal in this paper is to describe more precisely the long-time asymptotic behavior of a class of solutions initially close to a solitary wave. In the general study of the stability of periodic solutions of ordinary differential equations [CL], and traveling waves of parabolic systems [Sa], one often seeks to establish that a perturbed solution will approach, as $t \to \infty$, some fixed translate of the periodic orbit or wave. This is the property of *orbital asymptotic stability.* In the present context, since small perturbations of solitary waves can change the wave speed, what we seek to show is that under suitable conditions, if $u(x, t)$ is initially a small perturbation of a given solitary wave $u_c(x - ct + \gamma)$, then

$$u(x, t) - u_{c_+}(x - c_+ t + \gamma_+) \to 0 \quad \text{as } t \to +\infty \,, \tag{1.6}$$

for some c_+ near c and γ_+ near γ. If this property holds, we say that the *family* of solitary waves is asymptotically stable. (For the pure power nonlinearity with $f'(u) = u^p$, the family of solitary waves may also be regarded as a group orbit, under the larger group of symmetries consisting of translations $u \mapsto u(\cdot + \gamma)$, and dilations $u \mapsto c^{1/p} u(c^{1/2}\cdot)$.)

Now, the approach taken in the H^1 stability theory does not yield this information. The reason is as follows: To prove H^1 stability, the solitary wave profile, u_c, is viewed as a critical point of a conserved energy functional:

$$\mathscr{E}[u] = \mathscr{H}[u] + c\mathscr{N}[u] \,.$$

Here, the Hamiltonian $\mathscr{H}$ and impulse functional $\mathscr{N}$ are given by

$$\mathscr{H}[u] = \int \frac{1}{2}(\partial_x u)^2 - F(u)dx, \quad \mathscr{N}[u] = \int \frac{1}{2}u^2 dx \,, \tag{1.7}$$

where $F'(u) = f(u)$, $F(0) = 0$. The estimate (1.5) arises because u_c is a constrained minimum of $\mathscr{E}$, under the condition

$$d\mathscr{N}[u_c]/dc > 0 \,,$$

which is true for $p < 4$, due to the scaling relation $u_c(y) = c^{1/p} u_1(y\sqrt{c})$. Being derived from conserved integrals, the norm in (1.5) is insensitive to dispersive decay phenomena.

In order to establish asymptotic behavior of the type in (1.6), one should choose a norm which decreases as perturbations disperse. A program along these lines was carried out for a class of nonlinear Schrödinger equations by Soffer and Weinstein [SW1–3], who used L^p and polynomially weighted L^2 norms to establish the asymptotic stability of a family of nonlinear bound states. There, and in the present work, a key ingredient is a decay estimate for the *local energy* of perturbations, where the measure of local energy is tailored to the dynamics at hand.

What norm is appropriate for KdV solitons? For the KdV equation, numerical computations and certain results based on the method of inverse scattering suggest that, if a soliton moving to the right with speed $c > 0$ is perturbed, the solution will evolve toward a superposition of a similar dominant soliton, possibly followed by a number of small solitons (bound states) also propagating to the right, and then a dispersing part (radiation), cf. [ZK, GGKM, L, AS, DJ]. See Fig. 1. (For rigorous

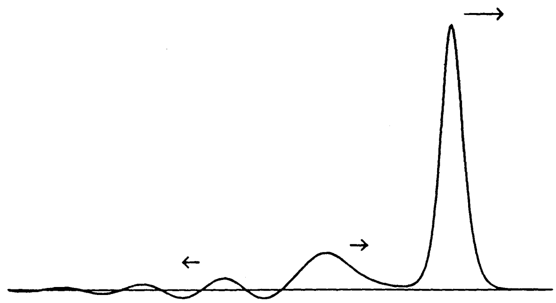

Fig. 1. Schematic picture of a solution of the KdV equation with two solitons and a dispersive wave

results regarding the emergence of solitons from arbitrary initial conditions, obtained using results from inverse scattering theory, see Schuur [Sc].)

One therefore should not expect (1.6) to occur in any translation-invariant norm, such as any L^p norm, $1 \leqq p \leqq \infty$. Heuristically, we may analyze the situation as follows: First, by (1.4), small solitary waves travel slower than larger ones; the amplitude increases with the speed c. Second, small amplitude dispersive waves, considered in the frame $y = x - ct$ of a dominant solitary wave traveling to the right with speed c, evolve approximately according to the equation

$$\partial_t u - c\partial_x u + \partial_x^3 u = 0\ ,$$

whose solutions are a superposition of plane waves $e^{ikx - i\omega t}$, where $\omega(k) = -ck - k^3$. The group velocity of these linear waves is always negative: $\omega'(k) = -c - 3k^2 < 0$ for all k. So small amplitude dispersive waves should travel to the left in this frame.

Hence, one may expect the dominant soliton to "outrun" the generated distortions. In a coordinate system moving with the large soliton, one could expect local uniform convergence in (1.6). Thus, we introduce a notion of local decay, to be used in the frame of the dominant solitary wave. This is expressed in terms of weighted norms, with exponential weights of the form e^{ay} where $a > 0$. We define

$$L_a^2 = \{v \mid e^{ay} v \in L^2(\mathbb{R})\}, \quad \text{with } \| v \|_{L_a^2} = \| e^{ay} v \|_{L^2}\ ,$$

$$H_a^1 = \{v \mid e^{ay} v \in H^1(\mathbb{R})\}, \quad \text{with } \| v \|_{H_a^1} = \| e^{ay} v \|_{H^1}\ .$$

Convergence in the space H_a^1 implies local uniform convergence. Furthermore, given a function $v(x + st)$ which is simply being translated to the left, with speed $-s < 0$, then its norm $\| v(\cdot + st) \|_{H_a^1}$ in the weighted space decays at an exponential rate, like e^{-ast}.

The global existence of solutions of the KdV equation with initial data $u(\cdot, 0) \in H^s \cap L_a^2$ with $s \geqq 2$ has been considered by Kato [K3], who showed that a unique solution exists with $u \in C([0, \infty), H^s \cap L_a^2)$, depends continuously on its initial data, and furthermore, enjoys a "parabolic" smoothing property, having $e^{ax} u \in C((0, \infty), H^{s'})$ for any real s'. Further developments concerning the well-posedness of KdV and gKdV appear in the more recent papers [KPV, GT].

Our main result for solitary waves of the KdV equation is as follows.

Theorem 1. *Let* $u_c(x - ct + \gamma)$, $c > 0$, $\gamma \in \mathbb{R}$, *be a solitary wave solution of the* KdV *equation* (1.1). *Suppose* $0 < a < \sqrt{c/3}$ *and* $0 < b < a(c - a^2)$. *Then there exists* $C > 0$ *such that if* $\varepsilon > 0$ *is sufficiently small, we have the following*: *Consider the initial value problem for the* KdV *equation with data*

$$u(x, 0) = u_c(x + \gamma) + v_0(x) . \tag{1.8}$$

Assume that $v_0 \in H^2 \cap H_a^1$, *with* $\| v_0 \|_{H^1} + \| v_0 \|_{H_a^1} < \varepsilon$. *Then there exist* $c_+ > 0$, $\gamma_+ \in \mathbb{R}$, *such that* $| c - c_+ | < C\varepsilon$, $| \gamma - \gamma_+ | < C\varepsilon$, *and for all* $t \geqq 0$ *we have*

$$\begin{aligned} &\| u(\cdot, t) - u_{c_+}(\cdot - c_+ t + \gamma_+) \|_{H^1} \leqq C\varepsilon, \\ &\| u(\cdot + c_+ t - \gamma_+, t) - u_{c_+}(\cdot) \|_{H_a^1} \leqq C\varepsilon e^{-bt} . \end{aligned} \tag{1.9}$$

Exactly the same result is true for the modified KdV *equation* (mKdV), *which is* (1.2) *with* $p = 2$.

Remarks.

1. The solution to KdV and gKdV (see Theorem 3 below) will be expressed in the form

$$u(x, t) = u_{c(t)}(x + \theta(t)) + v(x + \theta(t), t) , \tag{1.10}$$

where $\theta(t) = \gamma(t) - \int_0^t c(s)ds$. It is proved that

$$| c(t) - c_+ | + | \gamma(t) - \gamma_+ | + \| v(\cdot, t) \|_{H_a^1} \leqq C\varepsilon e^{-bt} .$$

The modulating speed $c(t)$ and phase $\gamma(t)$ do not depend on a, nor do their asymptotic limits c_+ and γ_+.

2. Related results for the KdV and mKdV ($p = 2$) equations appear in [Sc]. The KdV and mKdV equations are completely integrable, and may be solved by the inverse scattering transform. In [Sc] the representation of the solution in terms of the inverse scattering transform is analyzed to obtain information about the large time behavior of solutions in which solitons emerge. This approach does not apply to Eq. (1.2) with more general $f(u)$, where the equation is not expected to be integrable.

In (1.10), the leading (and dominant) term is an exact solitary wave solution of (1.1) when $c(t)$, $\gamma(t)$ do not vary in time. If we perturb the solitary wave slightly, it is natural to expect the solitary wave to adjust, via slow and small variations of its available parameters, to a nearby solitary wave. Thus we allow the parameters c and γ to "modulate." Substitution of the ansatz (1.10) into (1.2) yields an equation of the form

$$\partial_t v = \partial_y L_{c(t)} v - (\dot{c}\partial_c + \dot{\gamma}\partial_y) u_{c(t)} + \mathcal{S}(u_{c(t)}, v) , \tag{1.11}$$

where

$$L_c = -\partial_y^2 + c - f'(u_c) .$$

At this point $c(t)$ and $\gamma(t)$ are still unspecified functions of time. The evolution equations we obtain for these quantities may be said to arise from a *non-secularity condition* to be imposed on the solution v of (1.11). In the space L^2 (with domain

H^3), the operator $\partial_y L_c$ is degenerate, with an eigenvalue at the origin $\lambda = 0$. The (generalized) eigenfunctions are $\partial_y u_c$ and $\partial_c u_c$, which satisfy

$$\partial_y L_c \partial_y u_c = 0, \quad \partial_y L_c \partial_c u_c = -\partial_y u_c \, .$$

(To derive these equations, differentiate (1.3) with respect to y and c.) These two *zero modes* are associated, respectively, with infinitesimal changes in the location and speed of the solitary wave. They give rise to solutions $\partial_y u_c$ and $\partial_c u_c - t\partial_y u_c$ to the linearized problem

$$\partial_t v = \partial_y L_c v \, , \tag{1.12}$$

which are, respectively, constant and linearly growing with time.

Formally, to ensure that v contains no component of these solutions which exhibit secular growth, it is appropriate to require that the right-hand side of (1.11) be orthogonal to the (presumably 2-dimensional) generalized kernel of the adjoint of $\partial_y L_c$. These constraints yield two coupled first order differential equations for $c(t)$ and $\gamma(t)$ (called *modulation* equations), which are themselves coupled to the infinite dimensional dispersive evolution equation for $v(\cdot, t)$.

It turns out that the weighted space L^2_a also plays a role at this point. In fact, the generalized kernel of the adjoint of $\partial_y L_c$ is not 2-dimensional in the space L^2, but it is 2-dimensional in L^2_a for $0 < a < \sqrt{c}$. Thus, introducing the weighted space L^2_a provides a regularization which facilitates the derivation and justification of the modulation equations. (We note, however, that the modulation equations themselves do not depend on a.)

The functions $c(t)$ and $\gamma(t)$ are sometimes referred to as collective coordinates. Modulation equations for collective coordinates have been previously derived by various formalisms (see for example [KM, KA, Ne]). In formal perturbation theories, the coupling to the dispersion is usually neglected and the modulation equations are approximated by a coupled system of ordinary differential equations. The validity of this approximation on large but finite time intervals is considered in [W2] for a class of nonlinear Schrödinger equations.

Another point of view that describes our analysis is that the change of variables implicit in (1.10), from u to $(\gamma(t), c(t), v(y, t))$, is one for which the family of solitary waves becomes a 2-dimensional manifold of *equilibria*, corresponding to constant values of γ and c, with $v = 0$. We study the asymptotic stability of this manifold by regarding it as a center manifold. The "parabolic" character of the KdV equation in the space $H^s \cap L^2_a$ makes this approach feasible.

Our results below for gKdV in Theorem 3 will differ from the results in Theorem 1, due to differences arising in the detailed spectral properties of the operator $\partial_y L_c$ in the linearized evolution equation (1.12). As mentioned above, the point $\lambda = 0$ is an eigenvalue of the operator $\partial_y L_c$ in L^2. Concerning the rest of the spectrum, the results of [PW2] and Sect. 2 (see Theorem 2.1 below) imply that when $1 \leqq p < 4$, the spectrum consists of the entire imaginary axis. Most of this spectrum is approximate point spectrum. The point $\lambda = 0$ is an eigenvalue which is embedded in the essential spectrum.

A crucial spectral property that makes Theorem 1 possible is that for the solitary waves of KdV (and mKdV).

$$\lambda = 0 \text{ is the } \textit{only} \text{ eigenvalue of } \partial_y L_c \text{ in the space } L^2 \, . \tag{1.13}$$

In particular, the linearized equation (1.11) has no localized (L^2) solution of the form $e^{i\omega t}Y(y)$ with $\omega \neq 0$ real. While this refined spectral information concerning $\partial_y L_c$ is not required in the H^1-Lyapunov stability theory, it is necessary for our asymptotic stability analysis that (1.12) admit no spatially localized, temporally nondecaying solution which is not associated with modulation of the parameters $c(t)$ and $\gamma(t)$. We will prove (1.13) in sect. 3 and Appendix B. The proof relies on some general results in [PW2] concerning the eigenvalue problem for solitary waves of gKdV, and on explicit formulae available for the solution of the eigenvalue equation $\partial_y L_c v = \lambda v$ for KdV solitons. (Such formulae appear in [JK, Ber]. Our development relies on results from [M, GGKM].)

We are not presently able to prove that (1.13) holds for the solitary waves of gKdV for all $p \in (1, 4)$. For p fixed, we note that λ is an eigenvalue of $\partial_y L_c$ if and only if $\lambda/c^{3/2}$ is an eigenvalue of $\partial_y L_1$. (This is due to a dilational symmetry admitted by the gKdV equation, but can easily be checked using (1.4).) It follows that the property in (1.13) does not depend on c. What we can prove is the following:

Theorem 2. *The set* $\mathbf{E}$, *of values of* p *with* $p > 0$ *such that the operator* $\partial_y L_c$ *has a nonzero eigenvalue in* L^2, *is a discrete set. In particular,* $\mathbf{E} \cap [1, 4]$ *is a finite set (which does not contain the values* $p = 1$ *or* $p = 2$).

We conjecture that $\mathbf{E}$ is empty, in fact. There is strong numerical evidence to this effect, see the remarks concluding Sect. 3 below. But at this time, except for $p = 1$ and 2, we are unable to prove that any particular $p \in [1, 4]$ lies in $\mathbf{E}$ or not.

Our main stability result concerning gKdV solitary waves is as follows. We are interested in treating real values of p near $p = 4$, the transition to instability. Since for noninteger p, the nonlinearity $f(u)$ is not smooth, the results of Kato [K3] do not immediately yield the global existence of solutions. In Appendix A, we show that the method of Kato does yield global existence for $3 < p < 4$ (where f is C^4) : given $u(\cdot, 0) \in H^2 \cap H^1_a$, the solution $u \in C([0, \infty), H^2 \cap H^1_a)$, and $e^{ax}u \in C([0, \infty), H^{s'})$ for any $s' < 4$. In particular, the solution is *classical*: For $t > 0$, $\partial_t u$ and $\partial_x^3 u$ are continuous.

Theorem 3. *Let* $u_c(x - ct + \gamma)$, $c > 0$, $\gamma \in \mathbb{R}$, *be a solitary wave solution of the* gKdV *equation* (1.2). *Suppose* $3 \leqq p < 4$, *and assume that* (1.13) *holds,* i.e. $p \notin \mathbf{E}$. *Let* $0 < a < \sqrt{c/3}$. *Then there exists* $C > 0$ *and* b, $0 < b < a(c - a^2)$, *such that if* $\varepsilon > 0$ *is sufficiently small, we have the following: Consider the initial value problem for* gKdV *with the data in* (1.8). *Assume that* $v_0 \in H^2 \cap H^1_a$, *with* $\| v_0 \|_{H^1} + \| v_0 \|_{H^1_a} < \varepsilon$. *Then there exist* $c_+ > 0$, $\gamma_+ \in \mathbb{R}$, *such that* $| c - c_+ | < C\varepsilon$, $| \gamma - \gamma_+ | < C\varepsilon$, *and for all* $t \geqq 0$ *we have*

$$\begin{aligned} &\| u(\cdot, t) - u_{c_+}(\cdot - c_+ t + \gamma_+) \|_{H^1} \leqq C\varepsilon , \\ &\| u(\cdot + c_+ t - \gamma_+, t) - u_{c_+}(\cdot) \|_{H^1_a} \leqq C\varepsilon e^{-bt} . \end{aligned} \tag{1.14}$$

The conclusion of this theorem differs from that of Theorem 1 regarding the rate of exponential decay obtained in the weighted norm. In both Theorems 1 and 3, the local decay rate $-b$ satisfies $-a(c - a^2) < -b < 0$, but now it may be further restricted. The difference arises from the character of the spectrum of the operator $\partial_y L_c$ in (1.12), considered in the space L^2_a. Studying the resolvent equation

$(\lambda - \partial_y L_c)v = g$ in L^2_a, is equivalent, after multiplying by e^{ay} and letting $w = e^{ay}v$, $h = e^{ay}g$, to studying the resolvent equation

$$(\lambda - A_a)w = h, \quad \text{with } A_a = e^{ay}\partial_y L_c e^{-ay}, \tag{1.15}$$

in the space L^2. The transformation from $A_0 = \partial_y L_c$ to A_a has the effect of shifting the *essential spectrum* (defined to consist of all points of the spectrum which are not isolated eigenvalues of finite multiplicity [H, chap. 5]): The essential spectrum of A_0 is the imaginary axis, but the essential spectrum of A_a lies entirely in the left half plane $\operatorname{Re}\lambda \leqq -a(c-a^2) < 0$. Thus, modulo a finite dimensional subspace corresponding to point eigenvalues of A_a, we expect that the linearized flow defined by $e^{A_a t}$ is dissipative. Now, for the KdV and mKdV equations one can verify that the entire spectrum of A_a consists only of its essential spectrum, plus the isolated eigenvalue $\lambda = 0$ of algebraic multiplicity 2. But for gKdV this may no longer be true: In principle, the operator A_a can have isolated eigenvalues of finite multiplicity lying in the strip $-a(c-a^2) < \operatorname{Re}\lambda \leqq 0$. The property (1.13) is used to guarantee that, for $p \notin \mathbf{E}$, the only eigenvalue of A_a on the line $\operatorname{Re}\lambda = 0$ is $\lambda = 0$. The additional restriction on the exponential decay rate $-b$ in Theorem 3 arises because b must have the property that $\operatorname{Re}\lambda < -b$ whenever λ lies in the spectrum of A_a and $\lambda \neq 0$.

We remark that the restriction $0 < a < \sqrt{c/3}$ is imposed in Theorems 1 and 3 because the expression $a(c-a^2)$ is maximized at $a = \sqrt{c/3}$. Larger values of a would restrict the initial data further, with no gain in the decay rate achieved.

The possibility that A_a has isolated nonzero eigenvalues in the strip $-a(c-a^2) < \operatorname{Re}\lambda < 0$ becomes reality for p near 4, the point of transition to instability: For $p > 4$ the operator $A_0 = \partial_y L_c$ (and also A_a) has an eigenvalue $\lambda_\#(p) > 0$ [PW2–3]. Having characterized $\lambda_\#(p)$ as a zero of a Wronskian-like analytic function $D(\lambda)$ called Evans' function, Pego and Weinstein showed that $\lambda_\#(p)$ is analytic in p in a neighborhood of $p = 4$, with $\lambda_\#(p) < 0$ for $p < 4$.

When $\lambda_\#(p) < 0$, it is not an eigenvalue of A_0 in L^2. As discussed in [PW2–3], it is instead analogous to a *resonance pole* in quantum scattering theory [RS4]: It is a singularity arising in the analytic continuation of $(\lambda - A_0)^{-1}f(x)$ (for fixed x and $f \in L^2$ with compact support, for example), as λ moves from the right half plane, across the essential spectrum on the imaginary axis, onto the second sheet of a Riemann surface, above the left half plane. Such singularities of an analytically continued resolvent control radiation rates in a variety of physical problems, accounting, for example, for the phenomenon of Landau damping in the Vlasov–Poisson system of plasma physics [CH1–2], and for acoustic scattering in the wave equation, where local decay occurs at rates given by *scattering frequencies* [LP, V].

What we will show below is that in the present circumstances, $\lambda_\#(p)$ *is* a small negative eigenvalue of A_a, when $4 - p$ is small and positive. The decay rate b in Theorem 3 must then satisfy $\lambda_\#(p) < -b < 0$, and $\lambda_\#(p) \to 0$ as $p \to 4$. Therefore, the local decay rate of solitary wave perturbations, as guaranteed by Theorem 3, must approach zero as p approaches 4. See Fig. 2, comparing the spectrum of A_0 with that of A_a, for a value of p near 4.

Finally, a word about issues arising in carrying out *á priori* estimates of the perturbation about the solitary wave. A significant technical obstacle to overcome is that nonlinear terms like v^2 become discontinuous when considered as functions on the weighted space H^1_a, because the weight e^{ay} is not bounded away from zero.

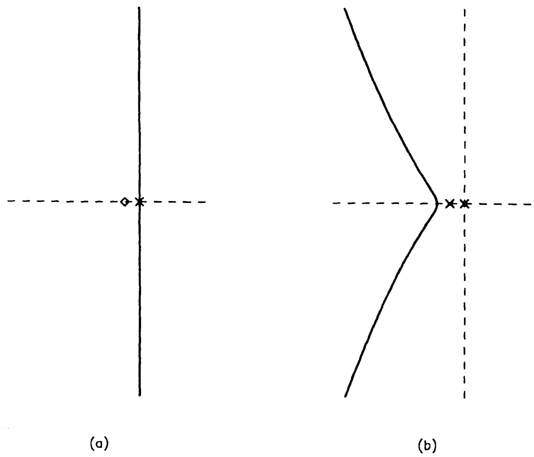

Fig. 2. (a) Spectrum of $\partial_y L_c$ for p near 4. '$\Diamond$' marks resonance pole. (b) Spectrum of A_a. '$\times$'s mark eigenvalues

(For this reason Sattinger [Sa] required his weights to remain strictly bounded away from zero.) As an example that describes how we will overcome this problem, we estimate $\| v^2 \|_{H^1_a}$ as follows:

$$\| v^2 \|_{H^1_a} = \| e^{ay} v^2 \|_{H^1} \leqq \| v \|_{H^1_a} \| v \|_{H^1} .$$

Thus if the unweighted norm $\| v \|_{H^1}$ can be shown to be small, then a quadratic term can be controlled like a small linear term. While control of the weighted perturbation $w(y, t) = e^{ay} v(y, t)$ is obtained by direct estimates of an integral equation, using smoothing and decay estimates on the semigroup $e^{A_a t}$, control of the unweighted norm $\| v(\cdot, t) \|_{H^1}$ requires a different kind of analysis. The key is to use the conserved energy functional $\mathscr{E}$, originating in the work of Benjamin [Be], for which u_c is a critical point. Using this together with the local decay of the perturbation $v(\cdot, t)$, we obtain the necessary bound for $\| v(\cdot, t) \|_{H^1}$.

The paper is organized as follows: In Sect. 2, we begin the spectral analysis of the linearized operator $\partial_y L_c$ in (1.12), characterizing its essential spectrum and generalized kernel in the spaces L^2 and L^2_a. Nonzero eigenvalues are characterized as zeros of Evans' function $D(\lambda)$, whose properties are recalled from [PW2] and further developed. In Sect. 3, we exhibit $D(\lambda)$ explicitly for $p = 1$ and 2, and verify the property (1.13). Also, we prove Theorem 2, by studying $D(\lambda, p)$ using analytic continuation in p.

We study the linear equation (1.12) in Sect. 4 by semigroup methods, and obtain certain smoothing and exponential decay estimates for later use. In Sect. 5 we justify the representation (1.10) of the solution, and derive the equations of motion of the new variables $(c(t), \gamma(t), w(y, t))$. In Sect. 6 we obtain the estimates indicated in Remark 1 following Theorem 1, and complete the proofs of Theorem 1 and 3.

Section 7 contains discussion of some further points, concerning, for example, multisoliton initial data, and the influence of the resonance pole on local asymptotic behavior when p is close to 4. In Appendix A, the existence and regularity of solutions of gKdV for $3 < p < 4$ is studied using Kato's theory. Appendix B contains the details of the calculation of Evans' function $D(\lambda)$ for the KdV and mKdV equations.

2. Spectral Analysis of the Linearized Equation

In this section we study the spectral properties of the operator $A_0 = \partial_y L_c$ appearing in the linearized equation (1.12), in the spaces L^2 and L^2_a.

2.1. Spectral Theory in L^2. We consider the operator $A_0 = \partial_y L_c$ on L^2 with domain H^3. The properties of the spectrum of this operator were delineated in [PW2]. The spectrum consists of *discrete* spectrum (isolated eigenvalues of finite multiplicity), and *essential* spectrum (everything else in the spectrum). Since $u_c(y) \to 0$ at an exponential rate as $|y| \to \infty$, the essential spectrum may be shown to agree with the spectrum S_e of the constant coefficient operator $\partial_y(-\partial_y^2 + c)$. Hence the essential spectrum S_e is the imaginary axis.

Regarding the isolated eigenvalues of A_0, the following result was proved in [PW2].

Theorem 2.1.

(1) *If $0 < p \leqq 4$ (corresponding to $d\mathcal{N}[u_c]/dc \geqq 0$), then A_0 has no isolated eigenvalues. Its spectrum coincides with the imaginary axis.*

(2) *If $p > 4$ (corresponding to $d\mathcal{N}[u_c]/dc < 0$), then the spectrum of A_0 consists of the imaginary axis together with two simple, real eigenvalues $\lambda = \lambda_\#(p) > 0$ and $-\lambda_\#(p) < 0$.*

In [PW2], the isolated eigenvalues of A_0 were studied using their characterization as zeros of Evans' function $D(\lambda)$. Evans' function also yields finer spectral information, such as the location of eigenvalues embedded in the essential spectrum, and resonance poles. This information is important in our asymptotic stability analysis.

We now discuss the definition of Evans' function $D(\lambda)$ and some of its key properties. For a more detailed development, see [PW2], also [E, AGJ]. If λ is an eigenvalue of A_0 with L^2-eigenfunction $Y(y)$, then Y is a solution of the differential equation

$$\partial_y[-\partial_y^2 + c - f'(u_c(y))]Y(y) = \lambda Y(y) . \tag{2.1}$$

As $|y| \to \infty$, the coefficients of (2.1) rapidly converge to those of the constant coefficient equation

$$\partial_y(-\partial_y^2 + c)Y(y) = \lambda Y(y) . \tag{2.2}$$

This equation has solutions of the form $e^{\mu y}$ where the exponent μ satisfies

$$\mathscr{P}(\mu) = -\mu^3 + c\mu = \lambda . \tag{2.3}$$

For arbitrary λ in the right half plane $\operatorname{Re}\lambda > 0$, Eq. (2.3) has roots $\mu_j(\lambda)$, $j = 1, 2, 3$, which satisfy

$$\operatorname{Re}\mu_1(\lambda) < 0 < \operatorname{Re}\mu_j(\lambda), \quad j = 2, 3 . \tag{2.4}$$

Corresponding to the solution $e^{\mu_1 y}$ of (2.2) which decays to zero as $y \to +\infty$, Eq. (2.1) has a solution $Y^+(y, \lambda)$ which is analytic in λ and satisfies

$$Y^+(y, \lambda) \sim e^{\mu_1 y} \quad \text{as } y \to +\infty . \tag{2.5}$$

From the solution $Y^+(y, \lambda)$, Evans' function $D(\lambda)$ may be defined, as a *transmission coefficient*, with the property that

$$Y^+(y, \lambda) \sim D(\lambda)e^{\mu_1 y} \quad \text{as } y \to -\infty\,. \tag{2.6}$$

$D(\lambda)$ is an analytic function in the right half plane. If $D(\lambda) = 0$ for some λ with $\mathrm{Re}\,\lambda > 0$, then $Y^+(y, \lambda)$ must decay exponentially as $y \to -\infty$. In this case, λ is an eigenvalue of A_0 with corresponding eigenfunction $Y = Y^+(\cdot, \lambda)$. Conversely, if λ is an eigenvalue with $\mathrm{Re}\,\lambda > 0$ and eigenfunction $Y(y)$, then $Y(y)$ must be a constant multiple of $Y^+(y, \lambda)$ and so $D(\lambda) = 0$, since $Y(y)$ is bounded.

Theorem 2.2. *For* $\mathrm{Re}\,\lambda > 0$, λ *is an eigenvalue of* A_0 *if and only if* $D(\lambda) = 0$.

The origin $\lambda = 0$ is an eigenvalue of A_0 embedded in the essential spectrum, with eigenfunction $\partial_y u_c$. Furthermore, $\partial_c u_c$ is a generalized eigenfunction: we have

$$\partial_y L_c \partial_y u_c = 0, \quad \partial_y L_c \partial_c u_c = -\partial_y u_c\,. \tag{2.7}$$

For the purposes of this paper, it is useful to observe that $D(\lambda)$ is naturally defined, by the same property (2.6), on a domain properly containing the (closed) right half plane, defined by the inequalities

$$\mathrm{Re}\,\mu_1(\lambda) < \mathrm{Re}\,\mu_j(\lambda), \quad j = 2, 3\,. \tag{2.8}$$

We denote the domain defined by (2.8) by Ω_0. For $\mathrm{Re}\,\lambda = 0$, it turns out that

$$\mathrm{Re}\,\mu_1(\lambda) < 0 = \mathrm{Re}\,\mu_2(\lambda) < \mathrm{Re}\,\mu_3(\lambda)\,,$$

thus (2.8) holds in a neighborhood of the imaginary axis, i.e., $\{\lambda \mid \mathrm{Re}\,\lambda \geqq 0\} \subset \Omega_0$. In fact, Ω_0 is explicitly given as follows.

Proposition 2.3. $D(\lambda)$ *is analytic in the whole complex plane, cut along the negative real axis from* $-\infty$ *to* $\lambda_* = -2(c/3)^{3/2}$. *That is,* $\Omega_0 = \mathbb{C}\backslash(-\infty, \lambda_*]$.

Proof. By the theory developed in [PW2], it suffices to show that the equation $\mathscr{P}(\mu) = \lambda$ has a unique root of smallest real part, for all $\lambda \in \mathbb{C}\backslash(-\infty, \lambda_*]$. This statement is true for $\mathrm{Re}\,\lambda \geqq 0$ because of (2.4), which is proved in [PW2]. Assume $\mathrm{Re}\,\lambda < 0$ and that $\mathscr{P}(\mu) = \lambda$ for distinct $\mu_1 = \alpha + i\beta_1$, $\mu_2 = \alpha + i\beta_2$, with the same real part. Now

$$\lambda = \mathscr{P}(\mu_j) = \alpha(c - \alpha^2) + 3\alpha\beta_j^2 + i\beta_j(c - 3\alpha^2 + \beta_j^2)\,.$$

We have $\alpha \neq 0$ since $\mathrm{Re}\,\lambda < 0$. Comparing real parts, we find $\beta_1^2 = \beta_2^2$, so $\beta_1 = -\beta_2$, hence $\mu_1 = \overline{\mu_2}$. Then comparing imaginary parts, we find that λ must be real, and $\beta_j^2 = 3\alpha^2 - c > 0$. Since $\lambda < 0$, we must have $\alpha < -\sqrt{c/3}$, so since $\lambda = \alpha(8\alpha^2 - 2c)$ is an increasing function of α, $\lambda < -\sqrt{c/3}(8c/3 - 2c) = \lambda_*$.

The only value of λ with $\mathrm{Re}\,\lambda < 0$ for which a double root occurs is when $\mathscr{P}'(\mu) = 0$. i.e., $\mu = -\sqrt{c/3}$ and $\lambda = \lambda_*$. The proposition now follows. □

In principle, when $\mathrm{Re}\,\lambda \leqq 0$, zeros of $D(\lambda)$ need not be eigenvalues of A_0, and conversely. However, the following was shown in [PW2], using the symmetry $Y(y) \mapsto \overline{Y(-y)}$ of (2.1) which is valid when $\mathrm{Re}\,\lambda = 0$.

Theorem 2.4. *Suppose* $\operatorname{Re}\lambda = 0$. *Then* λ *is an eigenvalue of* A_0 *if and only if* $D(\lambda) = 0$. *If* λ *is an eigenvalue, then the eigenfunction* $Y(y) = Y^+(y, \lambda) \to 0$ *at an exponential rate as* $|y| \to \infty$.

Since $\lambda = 0$ is an eigenvalue, $D(0) = 0$. It will turn out that $\lambda = 0$ is an isolated eigenvalue of A_0 in the space L^2_a. In order to describe the associated eigenspace and spectral projection, we introduce the following definitions:

$$\begin{aligned} \tilde{\xi}_1 &= \partial_y u_c, \quad \tilde{\xi}_2 = \partial_c u_c\,, \\ \tilde{\eta}_1 &= \theta_1 \int_{-\infty}^{y} \partial_c u_c + \theta_2 u_c, \quad \tilde{\eta}_2 = \theta_3 u_c\,. \end{aligned} \tag{2.9}$$

Here,

$$\theta_1 = \left(\frac{d}{dc}\mathcal{N}[u_c]\right)^{-1}, \ \theta_2 = \frac{1}{2}\left(\frac{d}{dc}\int_{-\infty}^{\infty} u_c\right)^2 \left(\frac{d}{dc}\mathcal{N}[u_c]\right)^{-2}, \text{ and } \theta_3 = -\theta_1\,.$$

The functions $\tilde{\xi}_1$ $\tilde{\xi}_2$, and $\tilde{\eta}_2$ decay exponentially as $|y| \to \infty$, at the rate $e^{-\sqrt{c}|y|}$. The function $\tilde{\eta}_1$ decays like $e^{\sqrt{c}y}$ as $y \to -\infty$, but is merely bounded as $y \to +\infty$. In addition, these functions have the following properties:

$$\begin{aligned} \partial_y L_c \tilde{\xi}_1 &= 0, \quad \partial_y L_c \tilde{\xi}_2 = -\tilde{\xi}_1\,, \\ L_c \partial_y \tilde{\eta}_1 &= \tilde{\eta}_2, \quad L_c \partial_y \tilde{\eta}_2 = 0\,, \end{aligned} \tag{2.10}$$

and

$$\langle \tilde{\eta}_j, \tilde{\xi}_k \rangle = \delta_{jk}, \quad j, k = 1, 2\,, \tag{2.11}$$

where $\langle u, v\rangle = \int_{-\infty}^{\infty} u\bar{v}\,dx$.

2.2. Spectral Theory in L^2_a. As mentioned in the remark following Theorem 1, we seek to prove that perturbations of a modulated solitary wave decay in a local energy sense, captured by norms in the weighted space L^2_a. Thus, we consider now the spectral theory of the linearized operator $\partial_y L_c$ in L^2_a.

We first make a change of variables,

$$W(y) = e^{ay}Y(y)\,. \tag{2.12}$$

Then the eigenvalue equation (2.1) is transformed into the equation

$$A_a W = e^{ay}\partial_y L_c e^{-ay} W = (\partial_y - a)[-(\partial_y - a)^2 + c - f'(u_c)]W = \lambda W\,. \tag{2.13}$$

The spectral theory of $A_0 = \partial_y L_c$ in L^2_a is equivalent to the spectral theory of A_a in L^2, and from now on we refer to the latter.

We first consider the essential spectrum of A_a. Since $f'(u_c(y))$ and $\partial_y f'(u_c(y))$ decay to zero at an exponential rate as $|y| \to \infty$, the essential spectrum of A_a can be shown to agree with the spectrum, S^a_e, of the constant coefficient operator

$$A^0_a = (\partial_y - a)[-(\partial_y - a)^2 + c]\,.$$

Hence, we have:

Proposition 2.5. *For* $0 < a < \sqrt{c/3}$, *the essential spectrum of* A_a *is the set* S_e^a, *a curve parametrized by*

$$\tau \mapsto \mathscr{P}(i\tau - a) = (i\tau - a)[-(i\tau - a)^2 + c]$$
$$= i\tau^3 - 3a\tau^2 + (c - 3a^2)i\tau - a(c - a^2)\,, \tag{2.14}$$

which lies in the open left half plane. (See Fig. 2.)

Next, we study the discrete spectrum of A_a. For $0 < a < \sqrt{c/3}$, the complement of the set S_e^a in the complex plane consists of two disjoint open components. One of these components, which we denote by Ω_+ or $\Omega_+(a)$, contains the closed right half plane. Any point of the spectrum of A_a lying in Ω_+ is an isolated eigenvalue of finite multiplicity. The following result characterizes such eigenvalues as zeros of Evans' function $D(\lambda)$.

Proposition 2.6. *Let* $0 < a < \sqrt{c/3}$. *Then:*

(i) $\{\lambda \mid \operatorname{Re}\lambda \geqq 0\} \subset \Omega_+(a) \subset \Omega_0$, *the domain of definition of* $D(\lambda)$.

(ii) *For* $\lambda \in \Omega_+(a)$, λ *is an eigenvalue of* A_a *in* L^2 *if and only if* $D(\lambda) = 0$.

Proof. Consider the curve in (2.14) which parametrizes S_e^a. The imaginary part of $\mathscr{P}(i\tau - a)$ is a strictly increasing function of τ, since $c - 3a^2 > 0$. Also $\mathscr{P}(-a) > \lambda_*$. So $S_e^a = \partial\Omega_+(a)$ and S_e^a does not intersect the cut $(-\infty, \lambda_*]$. Thus $\Omega_+(a) \subset \Omega_0$.

Next, observe that for $\lambda \in S_e^a$, the equation $\mathscr{P}(\mu) = \lambda$ has exactly *one* root satisfying $\operatorname{Re}\mu = -a$. Because of this, and the fact proved in [PW2] that the roots $\mu_j(\lambda)$ of (2.3) satisfy

$$\mu_j = (-\lambda)^{1/3} + O(|\lambda|^{-1/3}) \text{ as } |\lambda| \to \infty \text{ with } \lambda \in \Omega_0\,, \tag{2.15}$$

it follows that for $0 < a < \sqrt{c/3}$,

$$\operatorname{Re}\mu_1(\lambda) < -a = \operatorname{Re}\mu_2(\lambda) < \operatorname{Re}\mu_3(\lambda), \quad \lambda \in S_e^a\,,$$
$$\operatorname{Re}\mu_1(\lambda) < -a < \operatorname{Re}\mu_j(\lambda), \quad j = 2, 3, \quad \lambda \in \Omega_+(a)\,. \tag{2.16}$$

Now suppose $\lambda \in \Omega_+(a)$ is an eigenvalue of A_a. Then the differential equation (2.13) has solution $W(y)$ in L^2. By standard results on the asymptotic behavior of solutions of ordinary differential equations with asymptotically constant coefficients, $W(y)$ is bounded uniformly in y. Hence $e^{-ay}W(y)$ is a solution of (2.1) and satisfies

$$e^{-ay}W(y) = O(e^{-ay}) \text{ as } y \to \pm\infty\,.$$

Therefore, by Proposition 1.6 of [PW2], $e^{-ay}W(y)$ is a constant multiple of $Y^+(y, \lambda)$, and furthermore, since now $Y^+(y, \lambda) = O(e^{-ay})$ as $y \to -\infty$ and $\operatorname{Re}\mu_1(\lambda) < -a$ for $\lambda \in \Omega_+(a)$, it follows that $D(\lambda) = 0$. (Also see (2.6).)

If conversely, $D(\lambda) = 0$, then we know from Proposition 1.6 and Theorem 1.9 of [PW2] that

$$Y^+(y, \lambda) = \begin{cases} O(e^{\mu_1 y}) & \text{as } y \to +\infty\,, \\ O(e^{\mu_* y + \varepsilon|y|}) & \text{as } y \to -\infty\,, \end{cases}$$

whenever $0 < \varepsilon < \mu_* - \operatorname{Re}\mu_1$, where $\mu_* = \min(\operatorname{Re}\mu_2, \operatorname{Re}\mu_3)$. Hence by (2.16), $W(y) = e^{ay}Y^+(y, \lambda)$ satisfies (2.13) and decays exponentially as $|y| \to \infty$. So λ is an eigenvalue of A_a. $\square$

Proposition 2.6 implies in particular, since the closed right half plane $\{\lambda \mid \operatorname{Re}\lambda \geqq 0\} \subset \Omega_+$ for $0 < a < \sqrt{c/3}$, that eigenvalues of A_a in the closed right half plane must be zeros of $D(\lambda)$. We know that $\lambda = 0$ is an eigenvalue of A_0, and that $D(0) = 0$. Hence $\lambda = 0$ is an eigenvalue of A_a. In Sect. 3, we prove that for $1 \leqq p \leqq 4$, except for values $p \in \mathbf{E}$ (a finite set), $D(\lambda) \neq 0$ for all nonzero λ in the closed right half plane. (See Theorem 3.6.) Therefore, we have:

Theorem 2.7. *Let $0 < a < \sqrt{c/3}$ and assume $p = 1$ or 2, or $1 < p \leqq 4$ with $p \notin \mathbf{E}$. Then the only eigenvalue of A_a in the closed right half plane is $\lambda = 0$.*

We shall require a detailed characterization of the generalized eigenspaces of A_a and its adjoint $A_a^* = -e^{-ay} L_c \partial_y e^{ay}$. The dimension of these eigenspaces is determined by the fact that $0 = D(0) = D'(0) \neq D''(0)$ for $p \neq 4$, proved in [PW2]. (The fact that $0 = D(0) = D'(0)$ is associated with the existence of a two-parameter continuous family of solitary waves, obtained by translation and changes in wave speed.) For an operator A defined in L^2, define

$$\ker(A) = \{w \in \operatorname{dom}(A) \mid Aw = 0\}, \ \ker_g(A) = \bigcup_{k=1}^{\infty} \ker(A^k).$$

Proposition 2.8. (Spectral projections for the zero eigenvalue) *Assume $d\mathcal{N}[u_c]/dc \neq 0\ (p \neq 4)$ and $0 < a < \sqrt{c/3}$. Then $\lambda = 0$ is an eigenvalue for A_a with algebraic multiplicity two, and*

$$\ker_g(A_a) = \ker(A_a^2) = \operatorname{span}\{\xi_1, \xi_2\}, \ \ker_g(A_a^*) = \ker(A_a^{*2}) = \operatorname{span}\{\eta_1, \eta_2\}, \tag{2.17}$$

where $\xi_j = e^{ay}\tilde{\xi}_j$ and $\eta_j = e^{-ay}\tilde{\eta}_j$ for $j = 1, 2$, i.e.,

$$\begin{aligned} &\xi_1 = e^{ay}\partial_y u_c, \quad \xi_2 = e^{ay}\partial_c u_c, \\ &\eta_1 = e^{-ay}\left(\theta_1 \int_{-\infty}^{y} \partial_c u_c + \theta_2 u_c\right), \quad \eta_2 = e^{-ay}\theta_3 u_c, \end{aligned} \tag{2.18}$$

where $\theta_1, \theta_2, \theta_3$ are as in (2.9). *In addition, the ξ_j and η_k are biorthogonal, with $\langle \xi_j, \eta_k \rangle = \delta_{jk}$ for $j, k = 1, 2$. Thus the spectral projection P for A_a, associated with the eigenvalue $\lambda = 0$, and the complementary spectral projection Q, are given by*

$$Pw = \sum_{k=1}^{2} \langle w, \eta_k \rangle \xi_k, \quad Qw = (I - P)w = w - \sum_{k=1}^{2} \langle w, \eta_k \rangle \xi_k, \tag{2.19}$$

for $w \in L^2$. These projections satisfy $PA_a w = A_a P w$, $QA_a w = A_a Q w$, for $w \in \operatorname{dom}(A_a)$.

Proof. Because of (2.10), zero is an eigenvalue of A_a of algebraic multiplicity at least two, with

$$\begin{aligned} A_a\xi_1 = 0, \quad A_a\xi_2 = -\xi_1, \\ A_a^*\eta_1 = -\eta_2, \quad A_a^*\eta_2 = 0. \end{aligned} \tag{2.20}$$

We must show that the algebraic multiplicity is not greater than two; then the formula (2.19) for the spectral projection P follows. But there is a general relation between the algebraic multiplicity of eigenvalues of A_a and the order of vanishing of $D(\lambda)$, see [E], [AGJ]. We have the following.

Lemma 2.9. *Assume* $0 < a < \sqrt{c/3}$ *and* $\lambda \in \Omega_+(a)$ *with* $D(\lambda) = 0$. *The algebraic multiplicity of* λ *as an eigenvalue of* A_a *equals the order of* λ *as a zero of* $D(\lambda)$.

Proof. Assume λ is a zero of $D(\lambda)$ of order $k+1$, $k \geqq 0$, so $0 = D(\lambda) = \ldots = D^{(k)}(\lambda) \neq D^{(k+1)}(\lambda)$. By [PW2], Propositions 1.2 and 1.21, in this situation the derivatives $\partial_\lambda^j Y^+(y, \lambda)$ satisfy

$$\partial_\lambda^j Y^+(y, \lambda) = \begin{cases} O(e^{\mu_1 y + \varepsilon y}) & \text{as } y \to +\infty\,, \\ O(e^{\mu_* y + \varepsilon |y|}) & \text{as } y \to -\infty\,, \end{cases}$$

for any $\varepsilon > 0$, for $j = 0, \ldots, k$. Here $\mu_* = \min(\operatorname{Re}\mu_2, \operatorname{Re}\mu_3) > -a$ by (2.16). Define $W_j(y) = e^{ay}\partial_\lambda^j Y^+(y, \lambda)$. Since

$$(\partial_y L_c - \lambda)\,\partial_\lambda^j Y^+(y, \lambda) = j\partial_\lambda^{j-1} Y^+(y, \lambda)$$

for $j = 1, 2, \ldots$, we have that $W_j(y)$ decays exponentially as $|y| \to \infty$ for $j = 0, \ldots, k$, and

$$A_a W_0 = 0, \quad A_a W_j = jW_{j-1}, \quad \text{for } j = 1, \ldots, k\,.$$

Hence λ is an eigenvalue of A_a of algebraic multiplicity at least $k+1$; the functions $W_0, \ldots W_k$ form a Jordan chain.

To prove that the algebraic multiplicity is not greater than $k+1$, it remains to show:

(a) $\ker(A_a)$ is one dimensional, and (b) the equation $A_a W = W_k$ has no L^2 solution. That (a) is true is a consequence of the proof of Proposition 2.6. To prove (b), suppose that W is an L^2-solution of $A_a W = (k+1)W_k$. Then $W(y)$ is bounded. Put $Y(y) = e^{-ay}W(y) - \partial_\lambda^{(k+1)}Y^+(y, \lambda)$. We have $\partial_y L_c Y = 0$, $Y(y) = O(e^{-ay})$ as $y \to +\infty$. It follows that $Y(y)$ is a constant multiple of $Y^+(y, \lambda)$. Hence, we find that

$$\partial_y^j \partial_\lambda^{(k+1)} Y^+(y, \lambda) = O(e^{-ay}) \text{ as } y \to -\infty\,, \tag{2.21}$$

for $j = 0, 1, 2$. But Proposition 1.21 of [PW2] gives a formula that implies that since $D^{(k+1)}(\lambda) \neq 0$, the 3-vector $\partial_\lambda^{(k+1)}(Y^+, Y^{+'}, Y^{+''})$ has exact order $e^{\mu_1 y}$ as $y \to -\infty$. Since $\operatorname{Re}\mu_1 < -a$, the bound (2.21) contradicts this and implies $D^{(k+1)}(\lambda) = 0$. Hence (b) is true, and this finishes the proof of the lemma.

We conclude this section by describing how a resonance pole of the operator A_0, present for p near 4, yields an eigenvalue of A_a near 0. In [PW2] (see also [PW3]), we showed, for p less than but near the critical value $p_{\mathrm{cr}} = 4$, that $D(\lambda)$ has a real and negative zero, $\lambda_\#(p) < 0$, with $\lambda_\#(p) \to 0$ as $p \to 4$. Associated with $\lambda_\#(p)$ is the solution $\xi_\#(\cdot, p) = Y^+(\cdot, \lambda_\#(p))$ of (2.1), which for $p < 4$ decays exponentially as y tends to $+\infty$, and grows exponentially as $y \to -\infty$ since $\mu_*(\lambda_\#) < 0$:

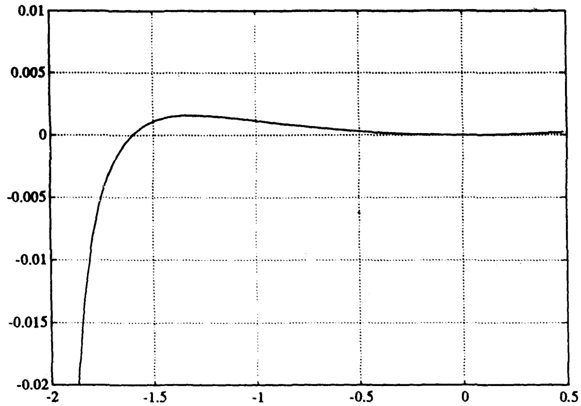

Fig. 3. $D(\lambda)$ vs. λ for $p = 3$, $c = 3$

$$\tilde{\xi}_{\#}(y, p) = \begin{cases} O(e^{\mu_1(\lambda_{\#})y}) & \text{as } y \to \infty\,, \\ O(e^{\mu_*(\lambda_{\#})y}) & \text{as } y \to -\infty\,. \end{cases} \tag{2.22}$$

The transition to instability of the solitary wave, as p increases from values less than $p_{\text{cr}} = 4$ to values larger than 4, is marked by the passage of $\lambda_{\#}(p)$ through the origin and becoming positive for $p > 4$. For $p > 4$, the function $\tilde{\xi}_{\#}(y, p)$ is now in L^2 since $\mu_*(\lambda_{\#}) > 0$, and so $\lambda_{\#}(p) > 0$ is now an unstable eigenvalue of A_0.

Now consider the function $\xi_{\#}(y, p) = e^{ay}\tilde{\xi}_{\#}(y, p)$. By (2.22) we have

$$\xi_{\#}(y, p) = \begin{cases} O(e^{(a+\mu_1(\lambda_{\#}))y}) & \text{as } y \to \infty\,, \\ O(e^{(a+\mu_*(\lambda_{\#}))y}) & \text{as } y \to -\infty\,. \end{cases} \tag{2.23}$$

Since $\lambda_{\#}(p) \to 0$ as $p \to 4$, we have $\mu_*(\lambda_{\#}(p)) \to 0$. Hence, for p sufficiently near 4 and $p < 4$, we have $\xi_{\#}(y, p) \in L^2$, and so $\lambda_{\#}(p) < 0$ is an eigenvalue of A_a. This eigenvalue (which is also a resonance pole of $\partial_y L_c$) is indicated in Fig. 2.

An analogous construction can be carried out for the adjoint operator $-L_c\partial_y$. Summarizing these results we have:

Proposition 2.10. *Let $0 < a < \sqrt{c}$. Let p be less than and sufficiently near $p_{\text{cr}} = 4$. Then*

(a) $\lambda_{\#}(p) < 0$ is an L^2 eigenvalue of A_a with corresponding exponentially decaying eigenfunction $\xi_{\#}(y, p)$.

(b) $\lambda_{\#}(p)$ is an L^2 eigenvalue of A_a^* with corresponding exponentially decaying eigenfunction $\eta_{\#}(y, p)$.

We can choose $\xi_{\#}(y, p)$ and $\eta_{\#}(y, p)$ to satisfy the normalization $\langle \xi_{\#}(p), \eta_{\#}(p)\rangle = 1$.

Remark . There is strong numerical evidence suggesting that no resonance pole is present for $1 \leqq p \leqq 2$ while there is a resonance pole for all p with $2 < p < 4$. In Fig. 3, we plot numerically computed values of $D(\lambda)$ vs. λ for $p = 3$ with $c = 3$, for λ between 0 and the endpoint of the cut $(-\infty, \lambda_*]$ where $\lambda_* = -2$, cf. Proposition 2.3. A zero, corresponding to a resonance pole of A_0, is apparent at the approximate value $\lambda_{\#}(3) \approx -1.6$.

3. Absence of Nonzero Embedded Eigenvalues

The results of this section will prove Theorem 2, establishing that for $p = 1$(KdV), $p = 2$ (mKdV), and for all but a possibly finite set of values of $p \in \mathbf{E}$ with $p \in (1,2) \cup (2,4)$, the linear operator $\partial_y L_c$ has no nonzero eigenvalues in L^2. By Theorems 2.3 and 2.4, this implies that $D(\lambda) \neq 0$ for nonzero λ with $\mathrm{Re}\lambda \geqq 0$, and hence that the operator A_a has no nonzero eigenvalues in the closed right half plane (except possibly when $p \in \mathbf{E}$). This proves Theorem 2.7. The proof of Theorem 2 relies on: (a) the explicit calculation of $D(\lambda)$ for the KdV and mKdV equations ($p = 1$ and 2), and (b) an analytic continuation argument in p, using a result of [PW2] which implies the *simplicity* of any nonzero embedded eigenvalue of $\partial_y L_c$. This argument does not depend on the explicit form of the nonlinearity $f(u, p)$, and could be used for other analytic families of nonlinearities that contain the KdV case $f(u) = u^2/2$.

Theorem 3.1. *Let $\lambda \in \Omega_0$ (see Proposition 2.3). (a) For the case of the KdV equation ($p = 1$, i.e., for the eigenvalue problem*

$$\partial_y L_c Y = \partial_y(-\partial_y^2 + c - 3c\,\mathrm{sech}^2(\frac{1}{2}y\sqrt{c}))Y = \lambda Y \ , \tag{3.1}$$

Evans' function is given explicitly by

$$D(\lambda) = \left(\frac{\mu_1(\lambda) + \sqrt{c}}{\mu_1(\lambda) - \sqrt{c}}\right)^2 , \tag{3.2}$$

where $\mu_1(\lambda)$ is the root μ of smallest real part of equation (2.3), $\mu^3 - c\mu + \lambda = 0$. *(b) For the case of the* mKdV *equation ($p = 2$), i.e, for the eigenvalue problem*

$$\partial_y L_c Y = \partial_y(-\partial_y^2 + c - 6c\,\mathrm{sech}^2(y\sqrt{c}))Y = \lambda Y \ , \tag{3.3}$$

Evans' function is also given explicitly by (3.2).

Corollary 3.2. *For the cases $p = 1$ and $p = 2$,*
(a) $\lambda = 0$ is the only eigenvalue of $\partial_y L_c$.
(b) $\lambda = 0$ is the only eigenvalue of A_a, for $0 < a < \sqrt{c/3}$.

To prove the corollary, note that if $D(\lambda) = 0$, then $\mu_1(\lambda) = -\sqrt{c}$, hence $\lambda = 0$ by (2.3). The proof of Theorem 3.1 is given in Appendix B.

Next, we study the eigenvalue problem for $\partial_y L_c$ for $p \notin \{1,2\}$:

$$\partial_y L_c Y = \partial_y(-\partial_y^2 + c - u_c^p(y))Y = \lambda Y \ , \tag{3.6}$$

where $u_c^p(y) = \frac{1}{2}c(p+1)(p+2)\,\mathrm{sech}^2(\frac{1}{2}yp\sqrt{c})$.

Lemma 3.3. *If λ is a nonzero purely imaginary eigenvalue of* (3.6), *then λ is simple and $D'(\lambda) \neq 0$.*

Proof. Recall from [PW2, Theorem 3.6] that if $\lambda = i\beta$ is an eigenvalue of (3.6) with $0 \neq \beta$ real, then the eigenfunction $Y(y) = Y^+(y, \lambda)$ decays exponentially as $y \to \pm\infty$. Moreover, the eigenspace is one dimensional (by [PW2, Proposition 1.6] any eigenfunction must be a multiple of $Y^+(y, \lambda)$). Furthermore, with $Y_\lambda^+(y) = \partial_\lambda Y^+(y, \lambda)$, the following ordinary differential equation is satisfied:

$$(\partial_y L_c - \lambda) Y_\lambda^+(y) = Y^+(y) . \tag{3.5}$$

Suppose $D'(\lambda) = 0$. Then from Propositions 1.2 and 1.21 of [PW2] we find that since $\mathrm{Re}\mu_2(\lambda) = 0$, for $k = 0, 1, 2$ we have

$$\partial_y^k Y_\lambda^+(y) \begin{cases} \to 0 \text{ exponentially as } y \to +\infty , \\ = o(e^{\varepsilon|y|}) \text{ as } y \to -\infty, \text{ for all } \varepsilon > 0 . \end{cases}$$

Therefore, when we multiply (3.5) by $\overline{L_c Y^+(y)}$ and integrate by parts, we are assured that the integrals involved converge and boundary terms vanish. We find

$$\begin{aligned} \langle Y^+, L_c Y^+ \rangle &= \int_{-\infty}^{\infty} Y^+ \overline{L_c Y^+} dy = \int_{-\infty}^{\infty} (\partial_y L_c - \lambda) Y_\lambda^+ \overline{L_c Y^+} dy \\ &= - \int_{-\infty}^{\infty} L_c Y_\lambda^+ \overline{\lambda Y^+} dy - \int_{-\infty}^{\infty} \lambda Y_\lambda^+ \overline{L_c Y^+} dy = 0 , \end{aligned}$$

since $\lambda + \bar{\lambda} = 0$ and L_c is real and formally self adjoint. It follows $\langle \overline{Y^+}, L_c \overline{Y^+} \rangle = 0$, and one has also

$$\lambda \langle Y^+, L_c \overline{Y^+} \rangle = \int_{-\infty}^{\infty} \partial_y L_c Y^+ L_c Y^+ dy = 0 .$$

From these considerations, we have that $\mathscr{Y} = \mathrm{span}\{Y^+(\cdot, \lambda), \overline{Y^+}(\cdot, \lambda)\}$ is a subspace of L^2 (complex valued functions) that satisfies

$$\langle L_c u, v \rangle = 0 \text{ for all } u, v \in \mathscr{Y} .$$

Furthermore, $\mathscr{Y} \cap \ker(L_c) = \{0\}$. Since L_c has only one negative eigenvalue, which is simple, it follows that $\dim \mathscr{Y} \leqq 1$. (This follows from Lemma 3.3 of [PW2], but is also easy to show directly.) But this contradicts the fact that $\dim \mathscr{Y} = 2$. Hence $D'(\lambda) \neq 0$.

We also claim that λ is simple, i.e., there is no L^2-solution of $(\partial_y L_c - \lambda)\tilde{Y} = Y$. If indeed there were, then from standard results on the asymptotic behavior of solutions of ordinary differential equations [CL, C], we find that $\tilde{Y}(y)$ (and its derivatives) decay to zero exponentially as $|y| \to \infty$. From this it follows that $Y_\lambda^+ - \tilde{Y}$ is a constant multiple of Y^+, and hence $Y_\lambda^+(y)$ (and its derivatives) decays exponentially as $|y| \to \infty$. But this implies that $D'(\lambda) = 0$; see Proposition 1.21 of [PW2]. This finishes the proof of Lemma 3.3. □

One consequence of Proposition 3.3 concerns how the zeros of $D(\lambda)$ depend on p. Considered as a function of λ and p, Evans' function $D(\lambda, p)$ is analytic in both λ and p, as remarked in [PW2]. So we have the following:

Corollary 3.4. *If for some positive β_0, p_0 it happens that $D(i\beta_0, p_0) = 0$, then there is an analytic function $\lambda_0(p)$, defined for p in some neighborhood of p_0, such that $\lambda_0(p_0) = i\beta_0$, and $D(\lambda_0(p), p) = 0$, and such that, for every (λ, p) in some small neighborhood of $(i\beta_0, p_0)$ with $D(\lambda, p) = 0$ one has $\lambda = \lambda_0(p)$.*

Our next result will be used to confine any zeros of $D(\lambda)$ to a compact set in the plane.

Lemma 3.5. *$D(\lambda) \to 1$ as $|\lambda| \to \infty$ with $\lambda \in \Omega_0$, uniformly for p in any compact set of $(0, \infty)$.*

Proof. It was proved in [PW2] that $D(\lambda) \to 1$ as $|\lambda| \to \infty$ with Re $\lambda > -\varepsilon$, for some $\varepsilon > 0$, for p fixed. But an examination of the proof, in Sects. 2b (iv) and 1g of [PW2], reveals that the stronger statement above is true, as a consequence of the following facts:

(1)
$$\int_{-\infty}^{\infty} |u_c^p| + |\partial_y(u_c^p)| dy$$

is uniformly bounded for p in any compact set of $(0, \infty)$, and also
(2) The roots $\mu_j(\lambda)$ of $\mathscr{P}(\mu) = \lambda$ in (2.3) do not depend on p, and satisfy (2.15), hence

$$|\mu_k^l / \mathscr{P}'(\mu_j)| = O(|\lambda|^{-1/3}) \text{ as } |\lambda| \to \infty \text{ with } \lambda \in \Omega_0 ,$$

for $l = 0, 1$ and any j, k.

The second fact was proved in [PW2], but the point is that it does not matter that $\mathrm{Re}(\mu_1 - \mu_2)$ can become arbitarily small for $\lambda \in \Omega_0$ as λ approaches the boundary. $\square$

Theorem 3.6. *The set* **E**, *of values of p in* $(0, \infty)$ *such that* $D(\lambda, p) = 0$ *for some nonzero imaginary* λ, *is a discrete set. In particular,* $\mathbf{E} \cap [1, 4]$ *is a finite set, which includes neither the value 1 nor the value 2.*

Proof. Assume that **E** has an accumulation point $p_0 > 0$. From Lemma 3.5, we may conclude that there is a real sequence $\beta_j > 0$ and distinct $p_j > 0$ such that $D(i\beta_j, p_j) = 0$, and that as $j \to \infty$, $p_j \to p_0$ and $\beta_j \to \beta_0 \geqq 0$.

We claim $\beta_0 > 0$. If instead $\beta_0 = 0$, a contradiction is obtained as follows. The value $\lambda = 0$ is known to be a zero of $D(\lambda)$ of order exactly *two* for all $p \neq 4$, *three* for $p = 4$ [PW2]. Since $D(\pm i\beta_i, p_j) = 0$ for all j, we infer that if $\beta_j \to 0$ then $\lambda = 0$ is a zero of $D(\lambda, p_0)$ of order at least four, contradicting known facts.

Now since $\beta_0 > 0$, we have $D(i\beta_0, p_0) = 0$. So from Corollary 3.4 there is an analytic curve $\lambda = \lambda_0(p)$ of zeros of $D(\lambda, p)$, defined for p near p_0, such that $i\beta_j = \lambda_0(p_j)$ for j sufficiently large. Since $p_j \to p_0$, we may be conclude that $\lambda_0(p)$ is purely imaginary for real p. (Consider the Taylor series of $i\lambda_0(p)$.) Hence the analytic continuation of $\lambda_0(p)$ will remain purely imaginary for all p in the maximal real interval of existence that contains p_0.

Now, the function $\lambda_0(p)$ may be analytically continued to be defined on the entire half line $p > 0$. This is a consequence of the results 3.3–3.5 above, the implicit function theorem, and the fact that $\mathrm{Im}\lambda_0(p) > 0$ for all p (which is proved as we proved $\beta_0 > 0$ above).

We conclude that for $p = 1$ in particular, $D(\lambda) = 0$ for $\lambda = \lambda_0(1)$, which is purely imaginary with positive imaginary part. But by inspection of the explicit formula (3.2), no such zero of $D(\lambda)$ exists for $p = 1$. This is a contradiction, and proves that **E** is discrete. We note that Theorem 2 is a corollary. $\square$

Remarks.

1. Although we have fixed the wave speed c in the above discussion, the set **E** does not change with c. This is due the the scaling satisfied by the the eigenvalue problem (3.6): λ is an eigenvalue (resp. zero of $D(\lambda)$) for $\partial_y L_c$ if and only if $c^{-3/2}\lambda$ is, for $\partial_y L_1$.

2. We cannot prove that $D(\lambda, p) \neq 0$ for purely imaginary $\lambda \neq 0$, for any particular $p > 0$, except $p = 1$ and $p = 2$. But for any particular value of p, strong numerical

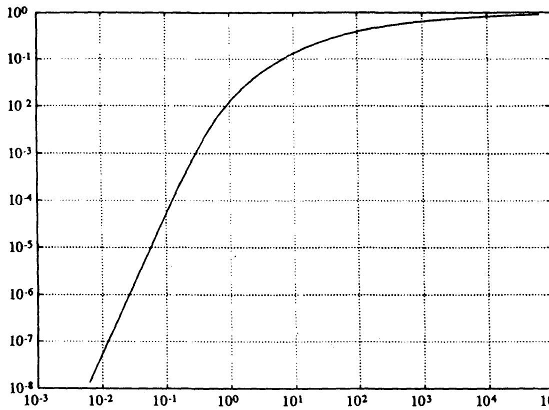

Fig. 4. Log-log plot of $|D(it)|$ vs. t, for $p = 4$

evidence is easy to obtain to decide whether it is so. For example, in Fig. 4, for the case $p = 4$ at which transition to instability occurs, we present a log-log plot of $|D(it)|$vs. log t. Note that for small t, the graph is approximately linear, with slope 3. Correspondingly, for $p = 4, \lambda = 0$ is known to be a zero of order 3 of $D(\lambda)$. The numerical evidence is strong for the following conjecture, which is significant for our study of the influence of a resonance pole on the decay rate of solitary wave perturbations when $4 - p > 0$ is small. See the remarks following Theorem 4.2 and the concluding Sect. 7 below.

Conjecture 3.7. *The value* $p = 4 \notin \mathbf{E}$. I.e., *for* $p = 4, D(\lambda) \neq 0$ *for imaginary* $\lambda \neq 0$.

4. Decay and Smoothing Estimates

In the introduction, we remarked that dispersing radiation, for the gKdV equation linearized about a constant state, moves to the left. One manifestation of this is that in a weighted space, with spatial weight decaying exponentially as $x \to -\infty$, the dynamics are *dissipative*. In this section, we develop the analysis of the linearized evolution equation (1.11), for solitary wave perturbations in such a space, the space L^2_a.

After the substitution

$$w(y,t) = e^{ay}v(y,t), \quad a > 0\,, \tag{4.1}$$

the linearized evolution equation (1.11) becomes

$$\partial_t w = A_a w, \text{ with } A_a = e^{ay}\partial_y L_c e^{-ay}\,. \tag{4.2}$$

As $|y| \to \infty$, the coefficients in (4.2) converge to those of the *free evolution equation*

$$\partial_t w = A_a^0 w\,, \tag{4.3}$$

where

$$A_a^0 = (\partial_y - a)(-(\partial_y - a)^2 + c) = -\partial_y^3 + 3a\partial_y^2 + (c - 3a^2)\partial_y - a(c - a^2)\,.$$

Indeed, $A_a = A_a^0 + (\partial_y - a)f'(u_c)$. Equation (4.2) is a *dissipative* evolution equation for $0 < a < \sqrt{c/3}$. As is easy to see using the Fourier transform, initial data in L^2 for (4.3) yield solutions which are C^∞ for $t > 0$ and decay exponentially to zero as $t \to \infty$:

Proposition 4.1. *For any integer $n \geqq 0$, and $0 < a < \sqrt{c/3}$, there exists $C = C(n,a)$ such that, for any $w \in L^2$ and for all $t > 0$,*

$$\|\partial_y^n e^{A_a^0 t} w\|_{L^2} \leqq C t^{-n/2} e^{-a(c-a^2)t} \|w\|_{L^2} .$$

Proof.

$$\partial_y^n e^{A_a^0 t} = e^{[-\partial_y^3 + (c-3a^2)\partial_y]t} e^{-a(c-a^2)t} \partial_y^n e^{3a\partial_y^2 t} .$$

The group $\exp\,(-\partial_y^3 + (c-3a^2)\partial_y)t$ is unitary on L^2. Therefore, we find that

$$e^{2a(c-a^2)t} \|\partial_y^n e^{A_a^0 t} w\|_{L^2}^2 = \|\partial_y^n e^{3a\partial_y^2 t} w\|_{L^2}^2 = \int_{-\infty}^{\infty} |\xi|^{2n} e^{-3a\xi^2 t} |\hat{w}(\xi)|^2 d\xi$$

$$\leqq \sup_\xi \left(|\xi|^{2n} e^{-3a\xi^2 t} \right) \|w\|_{L^2}^2 \leqq C t^{-n} \|w\|_{L^2}^2 .$$

The main result of this section is a decay and smoothing estimate for the semigroup $e^{A_a t}$, of the type above satisfied by the free semigroup $e^{A_a^0 t}$. Since $\lambda = 0$ is an eigenvalue of A_a, however, the estimate we seek will hold only on the invariant subspace range (Q) complementary to the generalized kernel of A_a.

Theorem 4.2. *Assume that $0 < a < \sqrt{c/3}$ and that $\lambda = 0$ is the only eigenvalue of A_a in the closed right half plane, with associated spectral projection P. Let $Q = I - P$. Then A_a is the generator of a C^0 semigroup on H^s for any real s, and, for any $b > 0$ such that the L^2-spectrum $\sigma(A_a) \subset \{\lambda | \mathrm{Re}\lambda < -b\} \cup \{0\}$, there exists C such that for all $w \in L^2$ and $t > 0$,*

$$\|e^{A_a t} Q w\|_{H_1} \leqq C t^{-1/2} e^{-bt} \|w\|_{L_2} . \tag{4.4}$$

Remark. The smoothing-decay estimate (4.4) will be used in the proofs of Theorem 1 (KdV and mKdV) and Theorem 3 (gKdV). For KdV and mKdV ($p = 1$ or 2), Corollary 3.2 (b) implies that for $0 < a < \sqrt{c/3}$, A_a has no eigenvalues in the open left half plane. Therefore, we can take $-b$, the exponentially rate of local energy decay, to satisfy $-a(c-a^2) < -b < 0$.

For general p, we can deduce from the results of Sect. 3 that if $p \notin \mathbf{E}$, then there is a number $b > 0$ for which the L^2-spectrum $\sigma(A_a) \subset \{\lambda | \mathrm{Re}\lambda < -b\} \cup \{0\}$. Furthermore, if $4 \notin \mathbf{E}$, and $4 - p > 0$ is sufficiently small, then Proposition 2.10 ensures the existence of a negative eigenvalue of $A_a, \lambda_\#(p)$, with $\lambda_\#(p) \to 0$ as $\to 4$. We then have for p less than and sufficiently near 4, that b is constrained by the inequality

$$\lambda_\#(p) < -b < 0 .$$

Therefore, the location of the resonance pole, $\lambda_\#(p)$, here dictates the exponential rate of decay of the perturbation's local energy.

The proof of Theorem 4.2 relies on perturbation arguments. The property that A_a generates a C^0 semigroup on H^s will be used below for $s = 0, 1$ and -3. To establish the estimate (4.4), we begin by studying the free resolvent $(\lambda I - A_a^0)^{-1}$, which is defined for λ not on the curve S_e^a in (2.14). It is convenient to estimate this resolvent in regions of the form $\overline{\Omega_+(\alpha)}$ for $0 < \alpha < a$; the boundary of $\Omega_+(\alpha)$ is the curve S_e^α, which will be a convenient contour to use for representing $e^{A_a t}Q$ as a contour integral.

Lemma 4.3. *Let $0 < \alpha < a < \sqrt{c/3}$. Then there exist C_0, C_1 such that for $\lambda \in \overline{\Omega_+(\alpha)}$ with $|\lambda| \geqq C_0$,*

$$\|\partial_y^n(\lambda I - A_a^0)^{-1}\| \leqq C_1|\lambda|^{(n-2)/3}, \quad \textit{for } n = 0, 1. \tag{4.5}$$

Here and below, $\|\cdot\|$ denotes the operator norm in L^2.

Proof. First, we note that the inequalities (2.16) imply that $\overline{\Omega_+(\alpha)} \subset \Omega_+(a)$ for $0 < \alpha < a < \sqrt{c/3}$. Also note that $(A_a^0 - \lambda I)e^{(\mu+a)y} = 0$ if and only if $\mathscr{P}(\mu) = \lambda$. The action of the resolvent is given by convolution with the Green's function $K_\lambda(y)$ for the resolvent equation $(\lambda I - A_a^0)v = w$, i.e., $(\lambda I - A_a^0)^{-1}w = K_\lambda * w$. Provided that $\mu_2(\lambda) \neq \mu_3(\lambda)$, we claim that the Green's function is given explicitly by

$$K_\lambda(y) = \begin{cases} a_1 e^{(\mu_1+a)y} & \text{for } y > 0, \\ -a_2 e^{(\mu_2+a)y} - a_3 e^{(\mu_3+a)y} & \text{for } y < 0, \end{cases} \tag{4.6}$$

where

$$a_j = a_j(\lambda) = 1/\prod_{k \neq j}(\mu_j - \mu_k).$$

To see this, we need to see that $(\lambda I - A_a^0)K_\lambda = \delta$, i.e., that with $[v] = v(0+) - v(0-)$, we have

$$[K_\lambda] = 0, \quad [K_\lambda'] = 0, \quad [K_\lambda''] = 1.$$

But this follows from a computation: Put $v_j = \mu_j + a$, then

$$\begin{aligned} a_1 + a_2 + a_3 &= \frac{(v_2 - v_3) - (v_1 - v_3) + (v_1 - v_2)}{(v_1 - v_2)(v_1 - v_3)(v_2 - v_3)} = 0, \\ v_1 a_1 + v_2 a_2 + v_3 a_3 &= \frac{v_1(v_2 - v_3) - v_2(v_1 - v_3) + v_3(v_1 - v_2)}{(v_1 - v_2)(v_1 - v_3)(v_2 - v_3)} = 0, \\ v_1^2 a_1 + v_2^2 a_2 + v_3^2 a_3 &= \frac{v_1^2(v_2 - v_3) - v_2^2(v_1 - v_3) + v_3^2(v_1 - v_2)}{(v_1 - v_2)(v_1 - v_3)(v_2 - v_3)} = 1, \end{aligned}$$

Now for λ large, it is true that $\mu_2 \neq \mu_3$ by (2.15). To prove the lemma, since $\partial_y^n(\lambda I - A_a^0)^{-1}w = \partial_y^n K_\lambda * w$, it suffices to estimate $\|\partial_y^n K_\lambda\|_{L^1}$ and use Young's inequality. From (4.6) we obtain the estimates

$$\|\partial_y^n K_\lambda\|_{L^1} \leqq \sum_j \frac{|\mu_j + a|^n}{|\mathrm{Re}\mu_j + a|} \prod_{k \neq j} \frac{1}{|\mu_j - \mu_k|}.$$

From (2.15) we have $1/|\mu_k - \mu_j| = O(|\lambda|^{-1/3})$ as $|\lambda| \to \infty$, and since $\lambda \in \overline{\Omega_+(\alpha)}$, from (2.16), $\mathrm{Re}\mu_j(\lambda) + a \geqq a - \alpha$ for $j = 2, 3$. From these facts and (2.15), one may also check that $\mathrm{Re}\mu_1 \to -\infty$ as $|\lambda| \to \infty$ with $\lambda \in \overline{\Omega_+(\alpha)}$. Hence we obtain the estimates

$$\|K_\lambda\|_{L^1} = O(|\lambda|^{-2/3}), \quad \|\partial_y K_\lambda\|_{L^1} = O(|\lambda|^{-1/3}),$$

as $|\lambda| \to \infty$ with $\lambda \in \overline{\Omega_+(\alpha)}$. The lemma follows.

Lemma 4.4 *Let* $0 < \alpha < a < \sqrt{c/3}$. *Then there exist* C_0, C_1 *such that for* $\lambda \in \overline{\Omega_+(\alpha)}$ *with* $|\lambda| \geqq C_0$, *we have* $\lambda \in \rho(A_a)$ *and*

$$\|\partial_y^n[(\lambda I - A_a)^{-1} - (\lambda I - A_a^0)^{-1}]\| \leqq C_1|\lambda|^{-1+n/3}, \quad n = 0, 1 . \tag{4.7}$$

$$\|\partial_y^n(\lambda I - A_a)^{-1}\| \leqq C_1|\lambda|^{(n-2)/3}, \quad n = 0, 1. \tag{4.8}$$

Proof. That $\lambda \in \rho(A_a)$ *if* $\lambda \in \overline{\Omega_+(\alpha)}$ is sufficiently large follows from Lemma 3.5 and Proposition 2.6. Now, if A and B are operators with the same domain, and $\lambda \in \rho(A) \cap \rho(B)$, then we have the resolvent identity

$$\begin{aligned}(\lambda I - B)^{-1} - (\lambda I - A)^{-1} &= (\lambda I - A)^{-1}(B - A)(\lambda I - A)^{-1} \\ &\quad \times [I - (B - A)(\lambda I - A)^{-1}]^{-1}.\end{aligned} \tag{4.9}$$

We may take $A = A_a^0, B = A_a$ in this identity, so $B - A = (\partial_y f'(u_c)) + f'(u_c)(\partial_y - a)$. Since $f'(u_c(y))$ and $\partial_y f'(u_c(y))$ are bounded, from Lemma 4.3 we have

$$\|(B - A)(\lambda I - A)^{-1}\| = \|(A_a - A_a^0)(\lambda I - A_a^0)^{-1}\| = O(|\lambda|^{-1/3})$$

as $|\lambda| \to \infty$ with $\lambda \in \overline{\Omega_+(\alpha)}$. From this we easily obtain (4.7), which together with Lemma 4.3 implies (4.8). □

Proof of Theorem 4.2. Postponing for a moment the proof that A_a is a generator of a C^0 semigroup on H^s, we complete the proof of (4.4). Let $b > 0$ be such that $\sigma(A_a) \subset \{\lambda | \mathrm{Re}\lambda < -b\} \cup \{0\}$. Then by Lemma 4.4, we may choose $\alpha, 0 < \alpha < a$, so that the nonzero spectrum of A_a lies to the left of the curve S_e^a, i.e.,$\overline{\Omega_+(\alpha)} \subset \rho(A_a) \cup \{0\}$. We may choose α so $-b < -\alpha(c - \alpha^2)$, so that the curve S_e^α intersects the line $\mathrm{Re}\lambda = -b$, at two points $\mathscr{P}(\pm i\tau_0 - \alpha)$ for some unique $\tau_0 > 0$ see (2.14).

We define the contour Γ to consist of the leftmost portions of the curve S_e^α and the vertical line $\mathrm{Re}\lambda = -b$. Γ may be parametrized by:

$$\tau \mapsto \lambda(\tau) = \begin{cases} \mathscr{P}(i\tau - \alpha) & \text{if } |\tau| \geqq \tau_0 , \\ -b + i\beta_0\tau & \text{if } |\tau| \leqq \tau_0 . \end{cases} \tag{4.10}$$

where $\beta_0 = \mathrm{Im}\mathscr{P}(i\tau_0 - \alpha)/\tau_0 > 0$.

Now, since $Q = I - P$, where P is the spectral projection for the eigenvalue $\lambda = 0$, the operator-valued function $\lambda \mapsto (\lambda I - A_a)^{-1}Q$ is analytic on and to the right of Γ, with only a removable singularity at $\lambda = 0$. Because of estimate (4.8), standard results in semigroup theory [P] imply that we have the representation

$$\begin{aligned} e^{A_a t}Q &= \frac{1}{2\pi i}\int_\Gamma e^{\lambda t}[(\lambda I - A_a)^{-1}Q d\lambda \\ &= e^{A_a^0 t}Q + \frac{1}{2\pi i}\int_\Gamma e^{\lambda t}[(\lambda I - A_a)^{-1} - (\lambda I - A_a^0)^{-1}]Q d\lambda . \end{aligned}$$

From Lemma 4.4, we obtain the following estimate for $n = 0, 1$. (Here, C denotes a generic constant, whose value may change from instance to instance.)

$$\|\int_\Gamma e^{\lambda t}\partial_y^n[(\lambda I - A_a)^{-1} - (\lambda I - A_a^0)^{-1}]Q d\lambda\| \leqq C\int_\Gamma e^{\operatorname{Re}\lambda t}|\lambda|^{-2/3}|d\lambda|$$

$$\leqq Ce^{-bt} + C\int_{\tau_0}^{\infty} e^{-\alpha(c-\alpha^2)t}e^{-3\alpha\tau^2 t}|\mathscr{P}(i\tau-\alpha)|^{-2/3}|\mathscr{P}'(i\tau-\alpha)|d\tau\,.$$

But $|\mathscr{P}'(i\tau-\alpha)||\mathscr{P}(i\tau-\alpha)|^{-2/3} \leqq C$, so the above is bounded by

$$Ce^{-bt} + Ce^{-\alpha(c-\alpha^2)t}\int_{\tau_0}^{\infty} e^{-3\alpha\tau^2 t}d\tau$$

$$\leqq Ce^{-bt} + Ce^{-\alpha(c-\alpha^2)t}e^{-3\alpha\tau_0^2 t}/\sqrt{t} = Ce^{-bt}(1+t^{-1/2})\,,$$

since $\operatorname{Re}\mathscr{P}(i\tau_0-\alpha) = -3\alpha\tau_0^2 - \alpha(c-\alpha^2) = -b$.

Now, in this argument, b can be replaced with a slightly larger value $b' > b$. So we may bound the above by

$$Ce^{-b't}(1+t^{-1/2}) \leqq C'e^{-bt}t^{-1/2}, \quad \text{for all } t > 0.$$

Combining this estimate with Lemma 4.1, the estimate (4.4) follows.

It remains to prove that in H^s, for any real s, A_a (with domain H^{s+3}) is the generator of a C^0 semigroup. Now, $A = A_a^0$ is the generator of a contraction semigroup on H^s; this is easy to check using the Fourier transform as in Proposition 4.1. We claim that, on H^s, the operator $B = A_a - A = (\partial_y - a)f'(u_c)$ has the following properties (the terminology is taken from Kato [K2]):

(i) B is A-bounded with relative bound 0, i.e., for any $\varepsilon > 0$,

$$\|Bu\|_{H^s} \leqq \varepsilon\|Au\|_{H^s} + C(\varepsilon)\|u\|_{H^s}, \quad u \in \operatorname{dom}(A); \tag{4.11}$$

(ii) B is quasi-accretive – it suffices to prove

$$|\langle Bu, u\rangle_{H^s}| \leqq C\|u\|_{H^s}^2, \quad u \in C_0^\infty. \tag{4.12}$$

By a standard result in perturbation theory [K2, p. 502], these properties imply that A_a is the generator of a C_0 semigroup on H^s. (Here, the inner product in H^s is given by $\langle u, v\rangle_{H^s} = \langle \Lambda^s u, \Lambda^s v\rangle$, where $\Lambda = (I - \partial_y^2)^{1/2}$.)

Property (i) is straightforward to prove, based on the two facts that: (a) $f'(u_c(\cdot))$ is smooth and all its derivatives decay exponentially, so it lies in H^s for all s; and (b) by standard interpolation estimates, for $j = 0, 1, 2$, the operator ∂_y^j is ∂_y^3-bounded with relative bound 0.

The proof of (ii) is based on a classic procedure of obtaining energy estimates and commutator estimates. Let $g(y) = f'(u_c(y))$, then clearly $\langle Bu, u\rangle_{H^s} = \langle g\partial_y u, u\rangle_{H^s} + O(\|u\|_{H^s}^2)$. Now with the notation $[A,B] = AB - BA$,

$$\langle \Lambda^s g\partial_y u, \Lambda^s u\rangle = \langle g\partial_y \Lambda^s u, \Lambda^s u\rangle + \langle [\Lambda^s, g\partial_y]u, \Lambda^s u\rangle\,.$$

The first term equals $-\frac{1}{2}\langle g'\Lambda^s u, \Lambda^s u\rangle = O(\|u\|_{H^s}^2)$. To bound the second term, we claim that $v = [\Lambda^s, g\partial_y]u$ satisfies $\|v\| = O(\|\Lambda^s u\|)$. To prove this, we write $\langle\xi\rangle = (1+\xi^2)^{1/2}$, then

$$\hat{v}(\xi) = \int_{-\infty}^{\infty} \hat{g}(\xi-\eta)(\langle\xi\rangle^s - \langle\eta\rangle^s)i\eta\hat{u}(\eta)d\eta = \int_{-\infty}^{\infty} K(\xi,\eta)\langle\eta\rangle^s\hat{u}(\eta)d\eta\,,$$

where $K(\xi,\eta) = \hat{g}(\xi-\eta)(\langle\xi\rangle^s - \langle\eta\rangle^s)i\eta/\langle\eta\rangle^s$. Now, to show that $\|v\|_{L^2} \leqq C\|\Lambda^s u\|_{L^2}$, it suffices to show that

$$\sup_{\xi} \int_{-\infty}^{\infty} |K(\xi,\eta)|d\eta + \sup_{\eta} \int_{-\infty}^{\infty} |K(\xi,\eta)|d\xi < \infty . \tag{4.13}$$

To estimate $|K|$, we make use of the inequalties $|\eta| \leqq \langle\eta\rangle$ and

$$\max(\langle\xi\rangle/\langle\eta\rangle, 1) \leqq |\langle\xi\rangle - \langle\eta\rangle| + 1 \leqq 2\langle\xi-\eta\rangle .$$

We consider two cases. If $s \geqq 1$, then

$$|\langle\xi\rangle^s - \langle\eta\rangle^s|/\langle\eta\rangle^{s-1} \leqq s\max((\langle\xi\rangle/\langle\eta\rangle, 1)^{s-1}|\xi-\eta| \leqq s2^{s-1}\langle\xi-\eta\rangle^s .$$

Hence $|K(\xi,\eta)| \leqq C|\hat{g}(\xi-\eta)|\langle\xi-\eta\rangle^s$. Since $\int_{-\infty}^{\infty} |\hat{g}(\xi)|\langle\xi\rangle^s d\xi \leqq C\|g\|_{H^{s+1}}$, (4.13) holds for $s \geqq 1$ If $s < 1$, then

$$\begin{aligned} |\langle\xi\rangle^s - \langle\eta\rangle^s|/\langle\eta\rangle^{s-1} &\leqq |s|\min(\langle\xi\rangle, \langle\eta\rangle)^{s-1}/\langle\eta\rangle^{s-1}|\xi-\eta \\ &= |s|\max(\langle\eta\rangle/\langle\xi\rangle, 1)^{1-s}|\xi-\eta| \leqq |s|2^{1-s}\langle\xi-\eta\rangle^{2-s}. \end{aligned}$$

From this, (4.13) similarly follows. □

Remark. It follows from the fact that A_a is a generator of a C^0 semigroup on H^1, that for $w \in H^1$,

$$\|e^{A_a t}w\|_{H^1} \leqq C\|w\|_{H^1}, \quad 0 \leqq t \leqq 1 .$$

Hence (4.4) implies also

$$\|e^{A_a t}Qw\|_{H^1} \leqq Ce^{-bt}\|w\|_{H^1}, \quad t \geqq 0 . \tag{4.14}$$

5. Decomposition of the Solution

We seek to represent solutions of the initial value problem (1.2), (1.8) for the gKdV equation in the form

$$u(x,t) = u_{c(t)}(y) + v(y,t) \tag{5.1}$$

with

$$y = y(x,t) = x - \int_0^t c(s)ds + \gamma(t) .$$

Given the initial data in (1.8), fix $c_0 = c$ and $\gamma_0 = \gamma$ to avoid a conflict of notation.

In order to achieve exponential decay for the perturbation $v(y,t)$ in the weighted space H_a^1, we wish to impose the constraint that

$$w(y,t) = e^{ay}v(y,t) \in \text{range}(Q) = \ker(P) , \tag{5.2}$$

where P is the spectral projection associated with the zero eigenvalue of the operator $A_a = e^{ay}\partial_y L_{c_0} e^{-ay}$. This requirement corresponds to the two scalar constraints $\langle w, \eta_k\rangle = 0, k = 1,2,$ cf.(2.19), which we will satisfy by modulating the parameters $\gamma(t), c(t)$ in a time-dependent fashion. An alternative point of view is that the change of variables $u(x,t) \mapsto (\gamma(t), c(t), v(y,t))$ is one for which the family of solitary waves becomes a manifold of equilibria (corresponding to γ, c constant, $v = 0$), and the representation (5.1) arises from the use of (time-dependent) tubular coordinates in a neighborhood of this manifold.

In this section, we first establish the existence of the decomposition (5.1) locally in time, for $u(\cdot,0)$ close to $u_{c_0}(\cdot+\gamma_0)$. We also establish a continuation property for this decomposition: it should persist as long as v remains small and $c(t)$ remains close to c_0. Finally, we establish the validity of evolution equations for $\gamma(t), c(t)$ and $v(y,t)$ which arise from the constraints on w. The *á priori* estimates carried out in Sect. 6 will be used to obtain the global continuation of the decomposition, and to prove the results on asymptotic behavior asserted in Theorems 1 and 3.

Proposition 5.1. (Local existence of the decomposition). *Let* $0 < a < \sqrt{c_0/3}$. *Let* s *be real, and* $t_1 \geqq 0$. *Then there exist* $\delta_0, \delta_1 > 0$ *such that: For any real* γ_0, *if* $u(x,t)$ *is such that*

$$e^{ax}u \in C([0,t_1],H^s) \textit{ with } \sup_{0\leqq t\leqq t_1} \|e^{a(\cdot+\gamma_0)}(u(\cdot,t)-u_{c_0}(\cdot-c_0t+\gamma_0))\|_{H_s} < \delta_0\,, \tag{5.3}$$

then there exists a unique function $t\mapsto(\gamma(t),c(t))$,

$$(\gamma,c)\in C([0,t_1],\mathbb{R}^2) \textit{ with } \sup_{0\leqq t\leqq t_1} |\gamma(t)-\gamma_0|+|c(t)-c_0| < \delta_1\,, \tag{5.4}$$

such that

$$\mathscr{T}_k[u,\gamma,c](t) =_{\text{def}} -\int_{-\infty}^{\infty}[u(x,t)-u_{c(t)}(y)]e^{ay}\eta_k(y)dx = 0\,, \tag{5.5}$$

for $k=1,2$, $0\leqq t\leqq t_1$, *where* $y = x-\int_0^t c(s)ds+\gamma(t)$. *The number* δ_0 *may be chosen as a decreasing function of* t_1. *The map* $u\mapsto(\gamma,c)$, *from the set defined in* (5.3) *to that defined in* (5.4), *is analytic, and moreover, if* $e^{ax}u\in C^m([0,t_1),H^s)$ *for some integer* $m>0$, *then* $(\gamma,c)\in C^m([0,t_1),\mathbb{R}^2)$.

Proof. The idea is to use the implicit function theorem [Ni] to solve (5.5) for the functions $(\gamma(t),c(t))$ in terms of u. First, from (1.4) and (2.18) we see that η_k and its spatial derivatives decay exponentially at infinity, so $\eta_k\in H^{-s}$ for all real s. In fact, one can check that $\eta_k(x)$ is analytic in a strip $|\operatorname{Im} x|<\varepsilon$ for some $\varepsilon>0$, so the map $\gamma\mapsto\eta_k(\cdot+\gamma)$ from $\mathbb{R}$ to H^{-s} is analytic. Also, the map $(\gamma,c)\mapsto u_c(\cdot+\gamma)$ from $\mathbb{R}\times\mathbb{R}^+$ to $H^s_a=\{v|e^{ax}v\in H^s\}$ is analytic. Then it is not hard to verify that the map $(u,\gamma,c)\mapsto\mathscr{T}=(\mathscr{T}_1,\mathscr{T}_2)$ defined in (5.5) is analytic, where $\mathscr{T}$ maps a neighborhood of the function $t\mapsto U_0(t)=(u_{c_0}(\cdot-c_0t),0,c_0)$ in $C([0,t_1],H^s_a)\times C([0,t_1],\mathbb{R}^2)$ to $C([0,t_1],\mathbb{R}^2)$. (Note that it suffices to consider the case $\gamma_0=0$, by a simple translation.)

In fact, $\mathscr{T}[U_0]=0$. To compute the Fréchet derivative of $\mathscr{T}$ with respect to the pair (γ,c) at U_0, observe

$$\mathscr{T}_k[u,\gamma,c](t) = -\int_{-\infty}^{\infty}u(x,t)\tilde\eta_k(y)dx+\int_{-\infty}^{\infty}u_{c(t)}(y)\tilde\eta_k(y)dy,$$

where $\tilde\eta_k, k=1,2$ are defined in (2.9). Then we find, for $(\delta\gamma,\delta c)\in C([0,t_1],\mathbb{R}^2)$, since $y=x-c_0t$ for $(u,\gamma,c)=U_0$,

$$\frac{\partial \mathcal{T}_k}{\partial \gamma}[U_0](\delta\gamma)(t) = -\int_{-\infty}^{\infty} u_{c_0}(x - c_0 t)\partial_y \tilde{\eta}_k(y)dx \cdot \delta\gamma(t)$$

$$= \int_{-\infty}^{\infty} \partial_y u_{c_0}(y)\tilde{\eta}_k(y)dy \cdot \delta\gamma(t) = \begin{cases} \delta\gamma(t) & \text{for } k = 1, \\ 0 & \text{for } k = 2, \end{cases}$$

$$\frac{\partial \mathcal{T}_k}{\partial c}[U_0](\delta c)(t) = \int_{-\infty}^{\infty} u(x,t)\partial_y \tilde{\eta}_k(y)dx \cdot \int_0^t \delta c(s)ds + \int_{-\infty}^{\infty} \partial_c u_{c_0}(y)\tilde{\eta}_k(y)dy \cdot \delta c(t)$$

$$= -\int_{-\infty}^{\infty} \partial_y u_{c_0}(y)\tilde{\eta}_k(y)dy \cdot \int_0^t \delta c(s)ds + \int_{-\infty}^{\infty} \partial_c u_{c_0}(y)\tilde{\eta}_k(y)dy \cdot \delta c(t)$$

$$= \begin{cases} -\int_0^t \delta c(s)ds & \text{for } k = 1, \\ \delta c(t) & \text{for } k = 2. \end{cases}$$

Hence the Fréchet derivative with respect to the pair (γ, c) may be written in block form as $\partial\mathcal{T}[u_0]/\partial(\gamma, c) = \begin{pmatrix} I & -B \\ 0 & I \end{pmatrix}$, where $(Bc)(t) = \int_0^t c(s)ds$ for $c \in C([0, t_1], \mathbb{R})$. This block operator is clearly invertible (replace $-B$ by B to get the inverse). Hence the implicit function theorem can be applied, proving the main conclusion of the Proposition, and establishing the analytic dependence of (γ, c) on u.

Since $\mathcal{T}[u, \gamma, c](t)$ depends only on values of $(u, \gamma, c)(s)$ for $0 \leqq s \leqq t$, one can follow the proof of the implicit function theorem to see that if a value δ_0 in (5.3) works for some value of $t_1 = t_1(\delta_0) > 0$, then it works for smaller values of t_1. So δ_0 may be chosen to increase as t_1 decreases.

Finally, if $u \in C^m([0, t_1), H_a^s)$, then if $0 < t_2 < t_1$, the curve $\tau \mapsto u(x, t + \tau)$ with values in $C([0, t_2], H_a^s)$ is a C^m curve for $0 \leqq \tau < t_1 - t_2$. Since the map $u \mapsto (\gamma, c)$ is analytic, and $u(\cdot, t + \tau) \mapsto (\gamma(t + \tau) - \int_0^\tau c(s)ds, c(t + \tau))$, it follows that $(\gamma, c) \in C^m([0, t_2], \mathbb{R}^2)$. □

Proposition 5.2. (Continuation principle). *There exist $\delta_0, \delta_1 > 0$ such that, for any $t > 0$, if*

$$e^{ax}u \in C([0, t_0], H^s) \quad \textit{with} \sup_{0 \leqq t \leqq t_0} \|e^{ay}v(\cdot, t)\|_{H^s} \leqq \delta_0/3\,, \tag{5.5}$$

where $v(y, t) = u(x, t) - u_{c(t)}(y), y = x - \int_0^t c(s)ds + \gamma(t)$, and if

$$(\gamma, c) \in C([0, t_0], \mathbb{R}^2) \quad \textit{with} \sup_{0 \leqq t \leqq t_0} |c(t) - c_0| \leqq \delta_1\,, \tag{5.6}$$

and $\mathcal{T}[u, \gamma, c](t) = 0$ for $0 \leqq t \leqq t_0$, then a unique extension of (γ, c) in $C([0, t_0 + t_*], \mathbb{R}^2)$ exists for some $t_* > 0$, with $\mathcal{T}[u, \gamma, c](t) = 0$ for $0 \leqq t \leqq t_0 + t_*$. Moreover, if $e^{ax}u \in C^m([0, \infty), H^s)$, then $(\gamma, c) \in C^m([0, t_0 + t_*), \mathbb{R}^2)$.

Proof. Let δ_0 be given by Proposition 5.1 for some $t_1 > 0$, and suppose δ_1 is such that $\|u_{c_1} - u_{c_0}\|_{H_a^s} \leqq \delta_0/3$ for $|c_1 - c_0| \leqq \delta_1$. Put $\tilde{u}(x, t) = u(x, t + t_0)$, and $\gamma_0 = -\int_0^{t_0} c(s)ds + \gamma(t_0)$. Then by (5.5) and the choice of δ_1 we have

$$\|e^{a(\cdot + \gamma_0)}(\tilde{u}(\cdot, 0) - u_{c_0}(\cdot + \gamma_0))\|_{H^s} \leqq$$
$$\|e^{ay}(u(\cdot, t_0) - u_{c(t_0)}(y))\|_{H^s} + \|u_{c(t_0)} - u_{c_0}\|_{H_a^s} \leqq 2\delta_0/3\,,$$

where $y = x - \int_0^{t_0} c(s)ds + \gamma(t_0) = x + \gamma_0$. It follows that $\tilde{u}$ satisfies the hypotheses of Proposition 5.1 for some sufficiently small $t_1 = t_* > 0$. From $\tilde{u}(x, t)$ one obtains $(\tilde{\gamma}(t), \tilde{c}(t))$, which must satisfy $(\tilde{\gamma}(0), \tilde{c}(0)) = (\gamma_0, c(t_0))$. Then the extension may be defined by $(\gamma(t), c(t)) = (\tilde{\gamma}(t-t_0) - \gamma_0 + \gamma(t_0), \tilde{c}(t-t_0))$ for $t_0 \leqq t \leqq t_0 + t_*$. The uniqueness and differentiability of (γ, c) may also be proved using the local result in Proposition 5.1. □

At this point, let us begin the proof of Theorem 1 and 3. Assume $p = 1, 2, 3$ or $3 < p < 4$. As discussed in Appendix A, the solution $u(x, t)$ of the initial value problem (1.2) – (1.8) satisfies, for any $T > 0$,

$$u \in C([0,T], H^2) \cap C^1([0,T], H^{-1}), \quad e^{ax}u \in C([0,T], H^1) \cap C^1([0,T], H^{-3}). \tag{5.7}$$

Moreover, u is a classical solution of (1.2) for $t > 0$. Given the initial data in (1.8), if $\|v_0\|_{H^1_a}$ is sufficiently small, it follows from Proposition 5.1 (taking $s = 1$ and $s = -3$) that the decomposition (5.1) exists locally in t, with $(\gamma, c) \in C^1([0, t_1), \mathbb{R}^2)$ for some $t_1 > 0$.

We now derive evolution equations for $\gamma(t), c(t)$, and $v(y, t)$, which are valid pointwise for $0 < t < t_1$. Substituting (5.1) into (1.2), we have

$$\begin{aligned} 0 &= \partial_t u + \partial_x^3 u + \partial_x f(u) \\ &= [\partial_t + (\dot{\gamma} - c(t))\partial_y + \partial_y^3](u_{c(t)} + v) + \partial_y(f(u_{c(t)} + v)) \\ &= [\partial_t - c_0\partial_y + \partial_y^3]v + \partial_y(f'(u_{c_0})v) \\ &\quad + (\dot{c}\partial_c + \dot{\gamma}\partial_y)u_{c(t)} + \partial_y[(\dot{\gamma} + c(t) - c_0)v + h(u_{c(t)}, v)v], \end{aligned}$$

where

$$h(u_{c(t)}, v)v = \int_0^1 [f'(u_{c(t)} + \tau v) - f'(u_{c_0})]d\tau\, v = f(u_{c(t)} + v) - f(u_{c(t)}) - f'(u_{c_0})v\,.$$

Thus $v(y, t)$ satisfies

$$\begin{aligned} \partial_t v &= \partial_y L_{c_0} v - (\dot{c}\partial_c + \dot{\gamma}\partial_y)u_{c(t)} \\ &\quad - \partial_y[(\dot{\gamma} + c(t) - c_0)v + h(u_{c(t)}, v)v]\,. \end{aligned}$$

Now $w(y, t) = e^{ay}v(y, t)$ satisfies (recall $A_a = e^{ay}\partial_y L_{c_0} e^{-ay}$)

$$\partial_t w = A_a w - \mathcal{F}, \tag{5.8}$$

where we write

$$\begin{aligned} \mathcal{F}(t) &= e^{ay}(\dot{c}\partial_c + \dot{\gamma}_y)u_{c(t)} + \dot{\gamma}e^{ay}\partial_y e^{-ay}w + \mathcal{G}(t)\,, \\ \mathcal{G}(t) &= e^{ay}\partial_y[c(t) - c_0 + h(u_{c(t)}, v)]e^{-ay}w\,. \end{aligned} \tag{5.9}$$

Equation (5.8) holds pointwise, but also in $C([0, t_1), H^{-3})$ due to (5.7). The constraint $w \in \text{range}(Q)$ in (5.2) now yields the following system of evolution equations for (w, γ, c), given v:

$$\partial_t w = A_a w + Q\mathcal{F}, \quad P\mathcal{F} = 0. \tag{5.10}$$

Written as an integral equation, the initial value problem for (5.10) becomes:

$$w(t) = e^{A_a t}w(0) + \int_0^t e^{A_a(t-s)}Q\mathcal{F}(s)ds\,. \tag{5.11}$$

This equation is initially justified in $C([0,t_1),H^{-3})$, but also holds in $C([0,t_1),L^2)$, since all terms lie in this space. The equation $P\mathscr{F}=0$ yields equations for $\dot\gamma,\dot c$ as follows. Introduce the notation

$$e_1(y,t)=e^{ay}(\partial_y u_{c(t)}(y)-\partial_y u_{c_0}(y))\,,$$

$$e_2(y,t)=e^{ay}(\partial_c u_{c(t)}(y)-\partial_c u_{c_0}(y))\,,$$

and note that $\langle(e^{ay}\partial_y e^{-ay}w,\eta_k\rangle=-\langle v,\partial_y\tilde\eta_k\rangle$ for $k=1,2$, by integration by parts. Then by (2.18), the condition $P\mathscr{F}=0$ is equivalent to

$$0=\langle\dot\gamma(\xi_1+e_1+(\partial_y-a)w)+\dot c(\xi_2+e_2)+\mathscr{G},\eta_k\rangle,\quad k=1,2\,. \tag{5.12}$$

Using the biorthogonality relations $\langle\xi_j,\eta_k\rangle=\delta_{jk}$, we obtain a system of equations for $\gamma(t)$ and $c(t)$:

$$\mathscr{A}(t)\begin{pmatrix}\dot\gamma\\ \dot c\end{pmatrix}=\begin{pmatrix}\langle\mathscr{G}, & \eta_1\rangle\\ \langle\mathscr{G}, & \eta_2\rangle\end{pmatrix},\mathscr{A}(t)=\begin{pmatrix}1+\langle e_1,\eta_1\rangle-\langle v,\partial_y\tilde\eta_1\rangle & \langle e_2,\eta_1\rangle\\ \langle e_1,\eta_2\rangle-\langle v,\partial_y\tilde\eta_2\rangle & 1+\langle e_2,\eta_2\rangle\end{pmatrix}. \tag{5.13}$$

The matrix $\mathscr{A}(t)$ satisfies

$$\mathscr{A}(t)=I+O(|c(t)-c_0|+\|v\|_{L^2})\quad\text{as}\quad|c(t)-c_0|+\|v\|_{L^2}\to 0\,. \tag{5.14}$$

Summarizing, we have that on some time interval $[0,t_1]$, the solution $u(x,t)$ of (1.2) can be decomposed as in (5.1) – (5.2), where $w(y,t)$, the weighted perturbation about the solitary wave, and $c(t),\gamma(t)$, the modulating speed and phase, satisfy the coupled system of equations (5.11), (5.13). Finally, from (5.12) or (5.13) it can be seen that the equations for $c(t)$ and $\gamma(t)$ do not depend on the weight parameter a; that is, (5.12) can be expressed entirely in terms of $c,\gamma,v(\cdot,t)$, and u_c.

6. Á Priori Estimates and Asymptotic Behavior

In this section we complete the proof of Theorems 1 and 3. What remains is to establish *á priori* estimates from the evolution equations in (5.11), (5.13). The *á priori* estimates will be seen to imply that the decomposition of solutions to (1.2), (5.1)–(5.2), persists for all time, with $v(y,\,t)$, the perturbation, remaining small in H^1 and decaying exponentially as $t\to+\infty$ in H^1_a.

Let $0<b<a(c-a^2)$ so the conditions of Theorem 4.2 are satisfied. (b is arbitrary for $p=1$ or 2, and in general, $\mathrm{Re}\lambda<-b$ for nonzero $\lambda\in\sigma(A_a)$.)

Proposition 6.1. *There exist $\delta_*>0$, $\varepsilon_0>0$, $C>0$ such that, if the decomposition* (5.1)–(5.2) *exists for $0\leqq t\leqq T$ and satisfies*

$$e^{bt}\|w(t)\|_{H^1}+|c(t)-c_0|+\|v(\cdot,\,t)\|_{H^1}\leqq\delta_*,\quad 0\leqq t\leqq T\,, \tag{6.1}$$

and if $\|v_0\|_{H^1}+\|v_0\|_{H^1_a}<\varepsilon\leqq\varepsilon_0$ in (1.8), *then*

$$e^{bt}\|w(t)\|_{H^1}+|c(t)-c_0|+\|v(\cdot,\,t)\|_{H^1}\leqq C\varepsilon,\quad 0\leqq t\leqq T\,. \tag{6.2}$$

Proof. The proof is broken down into two types of estimates:

(i) Local energy decay estimate, i.e. estimates of the weighted perturbation, $w(y,\,t)=e^{ay}v(y,\,t)$, in H^1, via the integral equation (5.11), and the modulation equations (5.12). Here we use the linear semigroup estimates of Theorem 4.2, which are valid provided $p \notin \mathbf{E}$, i.e., provided $\partial_y L_{c_0}$ has no nonzero eigenvalue.

(ii) H^1 estimate. Here, a Lyapunov functional (for which the solitary wave is a constrained minimum) is used, together with local decay estimates to control $\|v(\cdot, t)\|_{H^1}$.

(i) *Local Decay Estimate*: If δ_* is sufficiently small, then $\mathscr{A}(t)$ in (5.13) has bounded inverse, so we may estimate (5.13) to find

$$|\dot{\gamma}(t)| + |\dot{c}(t)| \leqq C\|\mathscr{G}(t)\|_{L^2} . \tag{6.3}$$

From (5.9), using that $e^{ay}\partial_y e^{-ay} = \partial_y - a$, we obtain the estimates

$$\|\mathscr{F}(t)\|_{L^2} \leqq C(|\dot{\gamma}|(1 + \|w\|_{H^1}) + |\dot{c}| + \|\mathscr{G}\|_{L^2}) \leqq C(1 + \|w\|_{H^1})\|\mathscr{G}\|_{L^2} ,$$

$$\|\mathscr{G}(t)\|_{L^2} \leqq C(|c(t) - c_0| + \|v\|_{H^1})\|w\|_{H^1} \leqq C\delta_*\|w\|_{H^1} .$$

Now, we may choose b' with $b < b' < a(c - a^2)$, such that b', as well as b, satisfies the conditions of Theorem 4.2. The remarks following the statement of Theorem 4.2 indicate how b is constrained for the case of KdV and gKdV. We may then estimate (5.11) as follows, for $0 \leqq t \leqq T$:

$$\begin{aligned} \|w(t)\|_{H^1} &\leqq Ce^{-b't}\|w(0)\|_{H^1} + C\int_0^t (t-s)^{-1/2}e^{-b'(t-s)}\|\mathscr{F}(s)\|_{L^2}\,ds \\ &\leqq Ce^{-b't}\|w(0)\|_{H^1} + C\int_0^t (t-s)^{-1/2}e^{-b'(t-s)}(1+\delta_*)\delta_*\|w(s)\|_{H^1}\,ds . \end{aligned} \tag{6.4}$$

Now define

$$M_w(T) = \sup_{0 \leqq t \leqq T} e^{bt}\|w(t)\|_{H^1} .$$

Then from (6.4) we find, for $0 \leqq t \leqq T$,

$$\begin{aligned} e^{bt}\|w(t)\|_{H^1} &\leqq C\|w(0)\|_{H^1} + C\delta_* M_w(T)\int_0^t (t-s)^{-1/2}e^{-(b'-b)(t-s)}\,ds \\ &\leqq C\|w(0)\|_{H^1} + C\delta_* M_w(T) . \end{aligned}$$

Taking the supremum over $0 \leqq t \leqq T$, we find that if δ_* is sufficiently small, then

$$M_w(T) = \sup_{0 \leqq t \leqq T} e^{bt}\|w(t)\|_{H^1} \leqq C\|w(0)\|_{H^1} . \tag{6.5}$$

Next, we estimate $|c(t) - c_0|$. Using (5.13) and (6.3) we find

$$\begin{aligned} |c(t) - c_0| &\leqq |c(0) - c_0| + \int_0^t |\dot{c}(s)|\,ds \leqq |c(0) - c_0| + \int_0^t C\delta_*\|w(s)\|_{H^1}\,ds \\ &\leqq |c(0) - c_0| + C\delta_* M_w(T) \leqq |c(0) - c_0| + C\delta_*\|w(0)\|_{H^1} . \end{aligned} \tag{6.6}$$

(ii) H^1 *Estimate*: We make use of the conserved quantity

$$\mathscr{E}[u] = \mathscr{H}[u] + c_0\mathscr{N}[u] = \int_{-\infty}^{\infty} \frac{1}{2}(\partial_x u)^2 - F(u) + \frac{1}{2}c_0 u^2\,dx ,$$

where $F'(u) = f(u)$, see (1.7). This is the Lyapunov functional used in the H^1 orbital stability results mentioned in the introduction. A key property of $\mathscr{E}$ is that u_c is a critical point. The key step in the Lyapunov stability analyses was to show that u_c is a local minimizer of $\mathscr{E}$ subject to the constraint of fixed L^2 norm. This requires a careful spectral analysis of the second variation of $\mathscr{E}$. Equipped with our local decay estimates $v = e^{-ay}w$, the H^1 estimate is now considerably simpler. Since u_{c_0} is a critical point of the functional $\mathscr{E}$, we have for any $z \in H^1$,

$$\mathscr{E}[u_{c_0} + z] - \mathscr{E}[u_{c_0}] = \int_{-\infty}^{\infty} \frac{1}{2}(\partial_x z)^2 + \frac{1}{2}(c_0 - f'(u_{c_0}))z^2 - g(u_{c_0}, z)z^3 \, dx \,, \tag{6.7}$$

where

$$\begin{aligned} g(u_{c_0}, z)z^3 &= \int_0^1 \frac{1}{2}(1-\tau)^2 f''(u_{c_0} + \tau z)d\tau z^3 \\ &= F(u_{c_0} + z) - F(u_{c_0}) - f(u_{c_0})z - \frac{1}{2}f'(u_{c_0})z^2 \,. \end{aligned}$$

Now, we take $z = u_{c(t)}(y) + v(y, t) - u_{c_0}(y) = u(x, t) - u_{c_0}(y)$ above, and observe that $\delta\mathscr{E}_0 = \mathscr{E}[u] - \mathscr{E}[u_{c_0}]$ is constant in time. We estimate (6.7) as follows. Note $\|u_{c(t)} - u_{c_0}\|_{H^1} \leqq C|c(t) - c_0|$ for δ_* sufficiently small. Then for some $k_1 > 0$,

$$\int_{-\infty}^{\infty} \frac{1}{2}(\partial_y z)^2 + \frac{1}{2}c_0 z^2 \, dy \geqq k_1 \|v\|_{H^1}^2 - C|c(t) - c_0|^2 \,.$$

Since $|f'(u_{c_0}(y))| \leqq C u_{c_0}(y)$ and $a < \sqrt{c}$, $e^{-ay}f'(u_{c_0}(y))$ is bounded in y. So we may estimate

$$\begin{aligned} \left| \int_{-\infty}^{\infty} f'(u_{c_0})z^2 \, dy \right| &\leqq \sup_y |e^{-ay}f'(u_{c_0}(y))| \, \|z\|_{L^2} \, \|e^{ay}z\|_{L^2} \\ &\leqq C(|c(t) - c_0| + \|v\|_{L^2})(|c(t) - c_0| + \|w\|_{L^2}) \\ &\leqq \frac{1}{4}k_1 \|v\|_{L^2}^2 + C(|c(t) - c_0|^2 + \|w\|_{L^2}^2) \,, \end{aligned}$$

where we have used the estimate $ab \leqq \delta a^2 + C(\delta)b^2$ for a suitably small δ. Finally, since $\|z\|_{H^1} \leqq C(|c(t) - c_0| + \|v\|_{H^1}) \leqq C\delta_*$,

$$\left| \int_{-\infty}^{\infty} g(u_{c_0}, z)z^3 \, dy \right| \leqq C\|z\|_{H^1}^3 \leqq C\delta_*(|c(t) - c_0|^2 + \|v\|_{H^1}^2) \,.$$

Hence, if δ_* is sufficiently small, (6.7) yields, with (6.5)–(6.6),

$$\frac{1}{2}k_1 \|v\|_{H^1}^2 \leqq \delta\mathscr{E}_0 + C(|c(t) - c_0|^2 + \|w\|_{L^2}^2) \leqq \delta\mathscr{E}_0 + C(|c(0) - c_0|^2 + \|w(0)\|_{H^1}^2) \,,$$

or

$$\|v\|_{H^1} \leqq C(\sqrt{\delta\mathscr{E}_0} + |c(0) - c_0| + \|w(0)\|_{H^1}) \,. \tag{6.8}$$

To finish the proof, it suffices to bound the right-hand side of (6.8) by $C(\|v_0\|_{H^1} + \|v_0\|_{H_a^1})$. Using (1.8) and $z = v_0$ in (6.7) we have

$$|\delta\mathscr{E}_0| \leqq C\|v_0\|_{H^1}^2 \,.$$

To estimate the remaining terms, observe from (1.8) and (5.1) that

$$u(x,0) = u_{c_0}(x+\gamma_0) + v_0(x) = u_{c(0)}(x+\gamma(0)) + v(x+\gamma(0),0)\,.$$

Applying Proposition 5.1 with $t_1 = 0$, we find that for $\|e^{ax}v_0\|_{H^1}$ sufficiently small, the map $u(\cdot,0) \mapsto (\gamma(0), c(0))$ is smooth on H^1_a, for $u(\cdot,0)$ near $u_{c_0}(\cdot+\gamma_0)$. In particular, this map is locally Lipschitz, and $u_{c_0}(\cdot+\gamma_0) \mapsto (\gamma_0, c_0)$, so

$$|\gamma(0)-\gamma_0| + |c(0)-c_0| \leqq C\|v_0\|_{H^1_a}\,.$$

From this it also follows

$$\|e^{ay}v(\cdot,0)\|_{H^1} = \|w(0)\|_{H^1} \leqq C\|v_0\|_{H^1_a}\,.$$

This completes the proof of Proposition 6.1. $\square$

We now complete the proof of Theorems 1 and 3. Let t_* be as in Proposition 5.2. First note that there is an $\varepsilon_1 > 0$, such that if $\|v_0\|_{H^1} + \|v_0\|_{H^1_a} < \varepsilon_1$, then for some $t_1 > 0$,

$$e^{ax}u \in C([0,t_1], H^1) \text{ with } \sup_{0\leqq t\leqq t_1} \|e^{a(\cdot+\gamma_0)}(u(\cdot,t) - u_{c_0}(\cdot - c_0 t + \gamma_0))\|_{H^1} < \delta_0\,.$$

Therefore, by Proposition 5.1, the decomposition (5.1)–(5.2) exists on the time interval $[0,t_1]$. Since the function $t \mapsto \|u(t)\|_{H^1} + \|u(t)\|_{H^1_a}$ is continuous, ε_1 can be chosen so that in addition (6.1) holds with $T = t_1$.

We denote by $T_{\max}$ the supremum of the set of all positive real numbers T, for which the solution $u(x,t)$ has a decomposition (5.1)–(5.2) for $t \in [0,T]$ and such that the estimate (6.1) holds. By the previous remark, $0 < T_{\max} \leqq \infty$. The proofs of Theorem 1 and Theorem 3 will be complete if we establish that $T_{\max} = +\infty$. If $T_{\max} < +\infty$, then we let $C\varepsilon_0 = \frac{1}{2}\min\{\delta_0/3, \delta_1, \delta_*, \varepsilon_1\}$, where δ_0, δ_1, δ_*, and C are as in Propositions 5.1, 5.2 and 6.1. Then, for $\|v_0\|_{H^1} + \|v_0\|_{H^1_a} < \varepsilon \leqq \varepsilon_0$, Proposition 6.1 implies that

$$e^{bt}\|w(t)\|_{H^1} + |c(t)-c_0| + \|v(\cdot,t)\|_{H^1} \leqq C\varepsilon_0, \quad 0 \leqq t \leqq T_{\max}\,. \tag{6.9}$$

The choice of ε_0 and Proposition 5.2 implies that the decomposition can be continued to yield a solution defined on the interval $[T_{\max}, T_{\max}+t_*]$. By the definition of ε_0, we have that the sum on the right-hand side of (6.9) is dominated by $\delta_*/2$. By continuity of $\|w(t)\|_{H^1}$, $c(t)$ and $\|v(\cdot,t)\|_{H^1}$, (6.1) holds with T replace by $T_{\max}+\tau$, for some τ with $0 < \tau$. This contradicts the definition of $T_{\max}$, and so $T_{\max} = +\infty$.

Continuing, we then have from (6.3), the estimate below it on $\|\mathscr{G}(t)\|_{L^2}$, and the bound on $\|w(t)\|_{H^1}$ implied by (6.9), that $|\dot{c}(t)| \leqq C\varepsilon e^{-bt}$. Hence $c_+ = \lim_{t\to\infty} c(t)$ exists, and $|c(t)-c_+| \leqq C\varepsilon e^{-bt}$. Similarly, $|\dot{\gamma}(t)| \leqq C\varepsilon e^{-bt}$, and so

$$\gamma_+ = \lim_{t\to\infty}\left(\gamma(t) - \int_0^t (c(s)-c_+)\,ds\right)$$

exists, and, defining $\tilde{\gamma}(t) = \gamma(t) - \int_0^t (c(s)-c_+)\,ds - \gamma_+$,

$$|\tilde{\gamma}(t)| = |\gamma(t) - \int_0^t c(s)\,ds + c_+ t - \gamma_+| \leqq C\varepsilon e^{-bt}\,.$$

We claim finally that the estimates (1.9) hold. Indeed,

$$u(x + c_+t - \gamma_+, t) - u_{c_+}(x) = u_{c(t)}(x + \tilde{\gamma}(t)) - u_{c_+}(x) + v(x + \tilde{\gamma}(t), t)\,.$$

Now we have the estimates

$$\|u_{c(t)}(\cdot + \tilde{\gamma}(t)) - u_{c_+}(\cdot)\|_{H^1 \cap H^1_a} \leqq C(|c(t) - c_+| + |\tilde{\gamma}(t)|) \leqq C\varepsilon e^{-bt}\,,$$
$$\|v(\cdot + \tilde{\gamma}(t), t)\|_{H^1_a} \leqq C\|w(\cdot\,, t)\|_{H^1} \leqq C\varepsilon e^{-bt}\,.$$

These estimates, together with the estimate $\|v(\cdot\,, t)\|_{H^1} \leqq C\varepsilon$ from (6.2), imply the estimates (1.9). This finishes the proof of Theorems 1 and 3. □

7. Further Discussion

In this paper we have studied solitary wave stability with respect to small perturbations in initial data, which are constrained to decay exponentially in space ahead of the wave, having small norm in $H^1 \cap H^1_a$. With such data, the solution asymptotically approaches a nearby solitary wave, at an exponential rate in the "local" sense implied by the norm in H^1_a. While the result in Theorem 1 is far less general than results achieved via inverse scattering by Schuur [Sc], concerning the emergence of any number of solitons from general initial data, the method involved makes minimal use of the Hamiltonian structure of the system, meaning that it may be more broadly useful.

One restriction that arises from our technical constraints on the initial perturbation concerns the possibility of small solitons emerging behind the main wave, in addition to any "dispersive radiation." In particular, consider the explicit N-soliton solutions of the KdV equation. In the large time limit, an N-soliton solution approaches a superposition of 1-solitons, arranged from left to right by order of increasing speed and amplitude, $c_1 < \cdots < c_N$. One may ask: if $N - 1$ of the wave speeds are sufficiently small, can the N-soliton be regarded as a small perturbation of a dominant 1-soliton to which Theorem 1 applies? The answer apparently must be *no*, for the following reason: If an N-soliton solution $u_N(x, t)$ corresponds asymptotically to 1-solitons with speeds $c_1 < \cdots < c_N$, then the *spatial* rate of decay of the solution is dictated by the amplitude of the smallest 1-soliton it contains: In fact, $u_N(x, t) \sim \alpha(t)e^{-\sqrt{c_1}x}$ as $x \to +\infty$ (cf. [GGKM2]). Thus, when we fix a in Theorem 1, this imposes a minimum size on the amplitude of the smallest wave in the combination, since we must have $\sqrt{c_1} > a$. But then there is no guarantee that the ordinary H^1 norm of the perturbation is small enough to meet the conditions of the theorem, since the constants involved do depend on a. We do, however, believe that by tracking the dependence of these constants on a, one ought to be able to improve the results to handle such initial data.

We have mentioned that the issue of existence of an asymptotic (or scattered) state is not addressed by the existing H^1 orbital stability theory, which asserts only that the solution remains close to some (time-varying) translate of the unperturbed wave. Recently, Bona and Soyeur [BSo] have improved this theory, for a large class of equations including (gKdV) and nonlinear Schrödinger equations, showing that the "wave speed" of the perturbed solution remains close to that of the unperturbed wave. For (gKdV), they identify this speed as the rate of change of the phase shift

$\gamma_u(t)$ which minimizes the L^2-distance from the solution to the solitary wave orbit, satisfying

$$\|u(\cdot, t) - u_c(\cdot - \gamma_u(t))\|_{L^2} = \inf_\gamma \|u(\cdot, t) - u_c(\cdot - \gamma)\|_{L^2} .$$

They do not show that the wave speed becomes asymptotically constant, but it may be that this is not true for all perturbations small in H^1. After all, such perturbations can decay slowly enough that the Schrödinger operator $-\partial_x^2 + u_0$ associated with the initial data can have an infinite number of discrete eigenvalues. In inverse scattering, this corresponds to the solution containing an infinite number of solitons. (The construction of such solutions has recently been announced by Gesztesy, Karwowski, and Zhao [GKZ].) Perhaps a 1-soliton may be perturbed by an "infinite" number of small solitons, so that the wave interactions cause the phase of the main wave to drift forever, and fail to converge.

Finally, we discuss two points connected with the decrease in the local decay rate e^{-bt} as guaranteed by Theorem 3, when p approaches 4 (assuming $4 \notin \mathbf{E}$, as is supported by numerical evidence, see Sect. 3). As discussed in the introduction, for gKdV this rate is constrained by the inequality

$$-a(c-a^2) < \lambda_\#(p) < -b < 0 ,$$

where $\lambda_\#(p)$ is a resonance pole of the operator $A_0 = \partial_y L_c$, and an eigenvalue of A_a, which approaches 0 as $p \to 4$. Indeed, in [PW2] it was proved that $\lambda_\#(p)$ is analytic in p near $p = p_{cr} = 4$, and has the following expansion, for some constant $\beta^2 > 0$:

$$\lambda_\#(p) = \beta^2 \cdot d\mathcal{N}[u_c]/dc + O((p - p_{cr})^2) .$$

(Here $d\mathcal{N}[u_c]/dc$ depends implicitly on p.) Thus, the quantity $d\mathcal{N}[u_c]/dc$, which arose in the H^1 orbital stability analysis as the quantity whose sign determines the stability of the wave, is also seen to have quantitative significance close to the transition to instability: It is proportional to the growth rate of the instability in the unstable regime $p > 4$, where $\lambda_\#(p) > 0$ is an eigenvalue of $\partial_y L_c$, and it is proportional to the maximum local decay rate in the stable regime $p < 4$.

Lastly, we believe that $\lambda_\#(p)$ does represent the true rate of local decay of the solitary wave perturbation when p is close to 4. Although the associated "eigenfunction" $\tilde{\xi}_\#(y)$ (described in Sect. 2) is not bounded as $y \to -\infty$, it is plausible that the "typical" perturbation may be asymptotically approximated, in the local sense of the norm in H_a^1, by an expression of the form

$$u(y + c_+ t - \gamma_+, t) - u_{c_+}(y) = \beta e^{\lambda_\# t} \tilde{\xi}_\#(y) + o(e^{\lambda_\# t}) \quad \text{as } t \to \infty .$$

While the profile of $\tilde{\xi}_\#(y)$ is unbounded, nevertheless we expect the shape of the perturbation to better approximate the shape of $\tilde{\xi}_\#(y)$ in the limit of large time on compact sets in y, as the amplitude decays exponentially to zero.

Appendix A. Existence Theory for Solutions

For the KdV equation, it is well known that the initial value problem is globally well posed in H^s for $s \geqq 2$, cf. [BSm, BSc, K3]. For gKdV with $f'(u) = u^p$, $p =$

1, 2, or 3, global well posedness in H^s, $s \geqq 2$, follows from the results of Kato [K3]. Kato also establishes well posedness in the space $H^s \cap L_a^2$, where L_a^2 is the exponentially weighted space $L_a^2 = \{w | e^{ax} w \in L^2\}$, $a > 0$. Kato showed that initial data in $H^s \cap L_a^2$ yield solutions of KdV which are C^∞ in x and t for $t > 0$. For recent results on the well posedness of gKdV, see [KPV].

For the results in this paper, we require some variations of earlier existence results. First, we must study solutions in $H^s \cap H_a^1$; this is only slightly different from the case $H^s \cap L_a^2$ studied by Kato. But also, we are interested in treating $f'(u) = u^p$ when $u > 0$, for noninteger values of p near 4, because of the influence of resonance poles as the transition to instability is approached. Because such nonlinearities are not smooth (f is only C^4 and not C^5 for $3 < p < 4$), some care ought to be taken to determine for which s one obtains an $H^s \cap H_a^1$ existence theory. We will obtain such a theory for $s = 2$ only, for $p > 3$ only; we collect the arguments necessary for this in this appendix.

Our strategy is to: (1) use Kato's abstract existence theory to establish local well posedness in H^2 for $f \in C^4$, in particular when $f'(u) = u^p$ for $u > 0$, where $p \geqq 3$ is real; (2) obtain global existence for $3 \leqq p < 4$ via standard *á priori* estimates (which are justified through regularization of f); and (3) remark that Kato's arguments in [K3] apply to obtain a solution $u \in C([0, \infty), H^2 \cap H_a^1)$, for $f' \in C^3$, $u_0 \in H^2 \cap H_a^1$.

(1) *Local well-posedness in* H^2. First, we recall Kato's abstract existence and well posedness theorem in the form we will use, cf. [K3, CS]: Consider an abstract quasilinear equation of evolution

$$\frac{du}{dt} + A(u)u = 0, \quad t \geqq 0, \quad u(0) = u_0 . \tag{A.1}$$

Let X, Y be real Hilbert spaces (with norms $\|\cdot\|_X$ and $\|\cdot\|_Y$ respectively), and assume $Y \subset X$ is dense with continuous injection. Assume $S : Y \mapsto X$ is an isomorphism of Y onto X.

(H1) Assume $A(y)$, defined for $y \in Y$, is a linear operator on X with domain $D(A(y)) \supset Y$, and $A(y)$ is quasi-m-accretive, uniformly for $\|y\|_Y$ bounded, i.e., given $R > 0$ there exists β such that for $\|y\|_Y \leqq R$,

$$\langle A(y)v, v\rangle_X \geqq -\beta \|v\|_X^2 \quad \text{for all } v \in D(A(y)) ,$$

and the range of $A(y) + \lambda$ is X for some (equivalently all) $\lambda > \beta$.

(H2) Assume that for any $R > 0$, there exists c_A such that

$$\|(A(y) - A(z))v\|_X \leqq c_A \|y - z\|_X \|v\|_Y \tag{A.2}$$

for all $y, z, v \in Y$ with $\|y\|_Y, \|z\|_Y \leqq R$.

(H3) For $y \in Y$, $SA(y)S^{-1} \supset A(y) + B(y)$, where $B(y) : X \mapsto X$ is a bounded linear operator, bounded uniformly for $\|y\|_Y$ bounded.

(H4) For any $R > 0$, there exists c_B such that

$$\|B(y)v - B(z)v\|_X \leqq c_B \|y - z\|_Y \|v\|_X \tag{A.3}$$

for all $y, z, v \in Y$ with $\|y\|_Y, \|z\|_Y \leqq R$.

The result of Kato is the following.

Theorem A.1. *Assume* (H1)–(H4). *Then for any* $R > 0$, *there exists* $T = T(R)$ *such that for any* $u_0 \in Y$ *with* $\|u_0\|_Y \leqq R$, *there exists a unique solution* u *of* (A.1) *with* $u \in C([0, T], Y) \cap C^1([0, T], X)$. *Also, the map* $u_0 \mapsto u$ *from* Y *to* $C([0, T], Y)$ *is continuous.*

In our application of this result to gKdV, we will follow Kato and take $Y = H^s$ with $s = 2$, $X = H^{s-3} = H^{-1}$, and $S = \Lambda^3$ where $\Lambda = (I - \partial_x^2)^{1/2}$. For $y \in Y$, $A(y)v$ is defined for $v \in Y = D(A(y))$ by

$$A(y)v = \partial_x^3 v + f'(y)\partial_x v \,. \tag{A.4}$$

We are interested in the case that f is C^r with $r \geqq 4$ only (corresponding to $f'(u) = u^p$ for $u > 0$ when $p \geqq 3$ is real), and proceed to verify hypotheses (H1)–(H4) in this case:

1. To verify (H1), we follow Kato [K1]. Since the operator ∂_x^3 in X with domain Y is skew-adjoint, by a perturbation theorem it suffices to show that for $\|y\|_Y \leqq R$, the operators $f'(y)\partial_x$ are uniformly quasi-accretive in X. It suffices to show

$$|\langle \Lambda^{-1} f'(y)\partial_x v, \Lambda^{-1} v\rangle| \leqq \beta \|\Lambda^{-1} v\|_{L^2}^2, \quad v \in H^2 \,, \tag{A.5}$$

for some β depending on R. Here $\langle \cdot, \cdot \rangle$ denotes the inner product on L^2 (real valued functions). To verify (A.5), let $z = \Lambda^{-1} v \in H^3$. Then, with the notation $[A, B] = AB - BA$, we have

$$\begin{aligned}\langle \Lambda^{-1} f'(y)\partial_x \Lambda z, z\rangle &= \langle \Lambda^{-1}[f'(y), \partial_x]\Lambda z, z\rangle + \langle \Lambda^{-1}\partial_x f'(y)\Lambda z, z\rangle \\ &= -\langle z, \Lambda(\partial_x f'(y))\Lambda^{-1} z\rangle - \langle [f'(y), \Lambda]z, \partial_x \Lambda^{-1} z\rangle \\ &\quad + \frac{1}{2}\langle (\partial_x f'(y))z, z\rangle \,.\end{aligned}$$

It is clear that the first and third terms are bounded by $\|\partial_x f'(y)\|_{H^1}\|z\|_{L^2}^2$. That the middle term is bounded by the same quantity is a consequence of the following lemma.

Lemma A.2. *Let* $b \in L^\infty$ *with* $\partial_x b \in H^1$. *Then*

$$\|[b, \Lambda]z\|_{L^2} \leqq C\|\partial_x b\|_{H^1}\|z\|_{L^2}, \quad z \in L^2 \,.$$

Proof. Let $\rho = [b, \Lambda]z$ and $\langle \xi \rangle = \sqrt{1 + \xi^2}$. Then

$$\hat{\rho}(\xi) = \int (\langle \eta \rangle - \langle \xi \rangle)\hat{b}(\xi - \eta)\hat{z}(\eta)d\eta \,.$$

Since $|\langle \eta \rangle - \langle \xi \rangle| \leqq |\xi - \eta|$ and $\| \, |\xi| \widehat{f' \circ g}(\xi)\|_{L_1} \leqq C\|\partial_x f'(y)\|_{H^1}$, by Young's inequality it follows that $\|\rho\|_{L^2} \leqq C\|\partial_x f'(y)\|_{H^1}\|z\|_{L^2}$.

To complete the proof of (A.5), observe that for $\|y\|_{H^2} \leqq R$, we have $\|D_x f'(y)\|_{H^1} \leqq C(R)$ since f' is C^2.

2. In order to verify (H2), it suffices to show

$$\|(f'(y) - f'(z))\partial_x v\|_{H^{-1}} \leqq c_A \|y - z\|_{H^{-1}}\|v\|_{H^2}, \quad v \in H^2 \tag{A.6}$$

for $\|y\|_{H^2}$, $\|z\|_{H^2} \leqq R$. We estimate the H^{-1} norm by duality: Let $h \in H^1$ with $\|h\|_{H^1} \leqq 1$. Define

$$r(y, z) = \begin{cases} \frac{f'(y)-f'(z)}{y-z} & \text{if } y \neq z, \\ f''(y) & \text{if } y = z. \end{cases}$$

Then

$$\begin{aligned} |\langle (f'(y) - f'(z))\partial_x v, h\rangle| &= |\langle y - z, r(y, z)(\partial_x v)h\rangle| \\ &\leqq \|y - z\|_{H^1} \|r(y, z)\partial_x v\|_{H^1} \\ &\leqq \|y - z\|_{H^1} \|\partial_x v\|_{H^1} (\|r(y, z)\|_{L^\infty} + \|\partial_x r(y, z)\|_{L^2}) . \end{aligned}$$

Now (A.6) follows from the following lemma.

Lemma A.3. *If f' is C^2, and $\|y\|_{H^1}$, $\|z\|_{H^1} \leqq R$, then*

$$\|r(y, z)\|_{L^\infty} + \|\partial_x r(y, z)\|_{L^2} \leqq C_R .$$

Proof. Use the representation $r(y, z) = \int_0^1 f''((1 - \tau)z + \tau y)\,d\tau$.

3. In order to verify (H3) it suffices to show that

$$\|[S, A(y)]z\|_X \leqq C_B \|z\|_Y, \quad z \in S^{-1}Y ,$$

since $S^{-1}Y = H^6$ is dense in Y. Thus we must show

$$\|[\Lambda^3, f'(y)]\,\partial_x z\|_{H^{-1}} \leqq C_B \|z\|_{H^2}, \quad z \in H^6 . \tag{A.7}$$

Since for f' in C^2, $\|y\|_{H^2} \leqq R$ implies $\|\partial_x f'(y)\|_{H^1} \leqq C_R$, (A.7) follows from this lemma:

Lemma A.4. *Let $b \in L^\infty$ with $\partial_x b \in H^1$. Then*

$$\|[\Lambda^3, b]\partial_x z\|_{H^{-1}} \leqq C \|\partial_x b\|_{H^1} \|\partial_x z\|_{H^1}, \quad z \in H^2 .$$

Proof. Let $h \in H^1$ with $\|h\|_{H^1} \leqq 1$. Then

$$\begin{aligned} |\langle [\Lambda^3, b]\partial_x z, h\rangle| &= |\langle (\Lambda[\Lambda^2, b] + \Lambda[b, \Lambda]\Lambda + [\Lambda^2, b]\Lambda)\partial_x z, h\rangle| \\ &\leqq \|[\Lambda^2, b]\partial_x z\|_{L^2} + \|[b, \Lambda]\Lambda\partial_x z\|_{L^2} + |\langle \Lambda\partial_x z, [\Lambda^2, b]h\rangle| \\ &\leqq C \|\partial_x b\|_{H^1} \|\partial_x z\|_{H^1} , \end{aligned}$$

where we have used Lemma A.2 to estimate the middle term.

4. In order to verify (H4), it suffices to check

$$\|[\Lambda^3, f'(y) - f'(z)]\partial_x z\|_{H^{-1}} \leqq C_B \|y - z\|_{H^2} \|z\|_{H^2} .$$

By Lemma A.4, this follows if we prove that for $\|y\|_{H^2}$, $\|z\|_{H^2} \leqq R$,

$$\|\partial_x (f'(y) - f'(z))\|_{H^1} \leqq C_R \|y - z\|_{H^2} .$$

But since f' is C^3, this is easily checked. (Note that when $f'(u) = u^p$ for $u > 0$ with $3 < p < 4$, f' is C^3 but not C^4.)

This completes the verification of hypotheses (H1)–(H4) of Theorem A.1, yielding local well posedness for gKdV in H^2 with f' in C^3.

(2) *Global existence in* H^2. In order to deduce from Theorem A.1 that solutions of (A.1) exist globally in t, when $f'(u) = u^p$ for $u > 0$ with $3 < p < 4$, it suffices to show that

$$\|u(t)\|_{H^2} \leqq C(T), \quad 0 \leqq t \leqq T, \tag{A.8}$$

for any $T > 0$. This *á priori* estimate follows from Kato's arguments in [K3], which are based on the following energy identities for solutions of (1.2):

$$\int u^2(x,t)\,dx = \int u_0^2(x)\,dx, \tag{A.9}$$

$$\mathscr{H}[u(t)] = \int \frac{1}{2}(\partial_x u)^2 - F(u)\,dx = \mathscr{H}[u_0], \tag{A.10}$$

$$\begin{aligned} \mathscr{E}_2[u(t)] &= \int (\partial_x^2 u)^2 - \frac{5}{3} f'(u)(\partial_x u)^2\,dx \\ &= \mathscr{E}_2[u_0] + \int_0^t \int \frac{1}{12} f^{(4)}(u)(\partial_x u)^5 + f'(u) f''(u)(\partial_x u)^3\,dx\,d\tau. \end{aligned} \tag{A.11}$$

We must establish the validity of these identities for the solutions u with $u \in C([0,T], H^2)$, given by Theorem A.1. Kato's strategy for the proof of the energy identities (A.9)–(A.11) is to regularize the initial data, and to work with solutions which have sufficiently many derivatives, so that the formal derivation of (A.9)–(A.11) is valid. We are unable to proceed directly in this manner due to the fact that f is only C^4. We therefore regularize the nonlinear function f as well. That is, we approximate (1.2) by a sequence of approximate problems

$$\begin{gathered} \partial_t u^n + \partial_x f^n(u^n) + \partial_x^3 u^n = 0, \\ u^n(x,0) = u_0^n(x), \end{gathered} \tag{A.12}$$

where $f^n \in C^\infty$ and $u_0^n \in H^s$ for all s, with $\|f^n - f\|_{C^4} \to 0$ and $\|u_0^n - u_0\|_{H^2} \to 0$ as $n \to \infty$.

From [K1, Theorem 7], it follows that the problems (A.12) are locally well posed in H^2 uniformly in n, in the following sense.

Theorem A.5. *There exists $T' > 0$ such that for all n, problem* (A.12) *has a unique solution $u^n \in C([0,T'], H^2)$. Moreover,*

$$u^n(t) \to u(t) \text{ in } H^2 \text{ uniformly for } t \in [0,T']. \tag{A.13}$$

In order to verify the hypotheses of Theorem 7 in [K1], it is necessary to verify that (H1)–(H4) hold uniformly in n (i.e., the hypotheses hold with constants not depending on n), and in addition:

$$\|(A^n(y) - A(y))v\|_X \to 0 \quad \text{for all } v \in Y, \tag{A.14}$$

$$\|(B^n(y) - B(y))v\|_X \to 0 \quad \text{for all } v \in X. \tag{A.15}$$

To prove (A.14) it suffices to check that

$$\|(f^{n\prime}(y) - f'(y))\partial_x v\|_{H^{-1}} \to 0 \text{ as } n \to \infty.$$

This follows easily by duality.

The proof of (A.15) is similar to the proof of (H4). One need only check that

$$\|\partial_x(f^{n\prime}(y) - f'(y))\|_{H^1} \to 0 \text{ as } n \to \infty, \text{ for } y \in H^2 .$$

But this is true because $f^n \to f$ in C^4.

In fact the approximating solutions u^n of (A.12) are smooth, with $u \in C([0, T'], H^s)$ for all s [K3, Theorem 4.1]; in particular T' is independent of s. It follows that the identities (A.9)–(A.11) hold for u^n, with F, f, replaced by F^n, f^n. We now let n tend to infinity, and use (A.13) to conclude that (A.9)–(A.11) hold for the limit, u.

From these considerations and Kato's arguments in [K3], the *á priori* estimate (A.8) follows, and global existence of solutions in H^2 is proved.

(3) *Solutions in* $H^2 \cap H^1_a$. Kato, in [K3], demonstrated that (1.2) is well posed on $H^s \cap L^2_a$ for any $s \geqq 2$, assuming f' is C^∞. Furthermore, a smoothing property holds: If $u_0 \in H^s \cap L^2_a$, then $u \in C([0, \infty), H^s \cap L^2_a)$, and $e^{ax}u \in C((0, \infty), H^{s'})$ for any $s' < s + 2$. These results were obtained by starting with the given solution $u \in C([0, \infty), H^s)$ and studying the properties of $w = e^{ax}u$, which is shown to lie in L^2 and satisfy the integral equation

$$w(t) = U_a(t)e^{ax}u_0 + \int_0^t U_a(t-r)f'(u(r))(\partial_x - a)w(r)dr , \tag{A.16}$$

where $U_a(t) = \exp(-(\partial_x - a)^3 t)$.

If we fix $s = 2$ and assume only $f' \in C^3$, Kato's arguments remain valid. Thus for $u_0 \in H^2 \cap L^2_a$, we obtain a solution

$$u \in C([0, \infty), H^2 \cap L^2_a), \quad \text{with } w = e^{ax}u \in C((0, \infty), H^{s'})$$

for all $s' < 4$. In particular, u is a *classical solution* of (1.2): Both $\partial_t u$ and $\partial_x^3 u$ are continuous in x, t for $t > 0$.

As a final remark, we note that if, in addition, $e^{ax}u_0 \in H^1$, then $w = e^{ax}u \in C([0, \infty), H^1)$. This may be proved from (A.16) using the smoothing properties of $U_a(t)$ and a bootstrap argument. Consequently, it is easy to see that the second term in (A.16) is in $C([0, \infty), L^2)$, and hence by standard arguments of semigroup theory [P], $w \in C^1([0, \infty), H^{-3})$.

Appendix B. Evans' Function for KdV and mKdV

Consider the Korteweg–de Vries equation

$$\partial_t u + u\partial_x u + \partial_x^3 u = 0 , \tag{KdV}$$

and the modified Korteweg–de Vries equation

$$\partial_t u + u^2\partial_x u + \partial_x^3 u = 0 , \tag{mKdV}$$

In this section we prove Theorem 3.1, which gives an explicit formula for Evans' function, $D(\lambda)$, associated with the linear eigenvalue problem

$$\partial_y L_c Y = \lambda Y, \quad L_c = -\partial_y^2 + c - u_c^p(y) , \tag{B.1}$$

in the cases $p = 1$ (KdV) and $p = 2$ (mKdV) (see (1.2)). Since $D(\lambda)$ is analytic, it suffices to establish the formula in Theorem 3.1 for $\operatorname{Re}\lambda > 0$, where we know that the roots of $\mu^3 - c\mu + \lambda = 0$ satisfy

$$\operatorname{Re}\mu_1(\lambda) < 0 < \operatorname{Re}\mu_j(\lambda), \quad j = 2, 3 .$$

The main tools used in this section are:

(i) the relationship between solution of the linearized KdV equation and the *Jost solutions* of an associated second order stationary Schrödinger equation [GGKM2],

(ii) the complex Miura transformation [M], which maps solutions of mKdV ($p = 2$) to solutions of KdV ($p = 1$), and

(iii) a trick of Darboux (1882), which gives a way of constructing solutions of a second order Schrödinger equation with potential $\tilde{u}(x)$, given solutions of some other such Schrödinger equation with potential $u(x)$. (See [T, Ke], for example. We thank R. Krasny for pointing this tool out to us.)

In order to construct $D(\lambda)$ we must study the behavior, as $y \to -\infty$, of $Y^+(y, \lambda)$, the solution of Eq. (2.1) with *maximal decay* as $y \to \infty$ (see Sect. 2). Our object in this section is to show that for KdV ($p = 1$) and mKdV ($p = 2$), the construction of $Y^+(y, \lambda)$ can be reduced to finding the solution of a second order Schrödinger equation,

$$-\partial_y^2 f(y, k) + (-\frac{1}{6}z(y) - k^2) f(y, k) = 0 , \tag{B.2}$$

with the asymptotic behavior

$$e^{-iky} f(y, k) \to 1 \text{ as } y \to +\infty . \tag{B.3}$$

A solution of (B.2) satisfying the asymptotic condition (B.3) is called a Jost solution [RS3].

Lemma B.1. *Let $z(x - ct)$ be a traveling wave solution of KdV, i.e. z satisfies $z(y) \to 0$ as $y \to \pm\infty$ and the second order nonlinear ordinary differential equation*

$$-cz + \frac{1}{2}z^2 - z'' = 0 . \tag{B.4}$$

Let $f(y, k)$ denote a (Jost) solution of the Schrödinger equation (B.2) *with the asymptotic behavior* (B.3). *Let $\mu = 2ik$ and $\lambda = c\mu - \mu^3$, and suppose that $Re\lambda > 0 > Re\mu$. Then,*

$$(-\partial_y^2 + c - z)\partial_y V = \lambda V , \tag{B.5}$$

where $V(y, k) = f(y, k)^2$.

Remark . This proposition allows z to be real or complex valued. Our application to KdV will involve choosing z equal to the real solitary wave profile u_c. For mKdV we shall apply Proposition B.1 with z given by a complex solitary wave of KdV.

Proof. An explicit calculation gives

$$(-\partial_y^2 - \frac{1}{6}z - k^2)\left[\frac{1}{2f}(-\partial_y^2 + c - z)\partial_y V\right] = 0 .$$

Therefore

$$\phi(y) \equiv \frac{1}{f}(-\partial_y^2 + c - z)\partial_y V = af(y, k) + bg(y) ,$$

where a and b are constants, and g is a solution of (B.2) which is linearly independent of f. Now, by the theory of asymptotic behavior of solutions of ordinary differential equations, $z(y) \to 0$ as $y \to \pm\infty$, and $g(y)$ is unbounded, being asymptotically proportional to e^{-iky} as $y \to +\infty$. Since it is also true that

$$\partial_y^j f(y, k) \sim (ik)^j e^{iky} \text{ as } y \to +\infty \text{ for } j = 0, 1, 2, 3 ,$$

we find that

$$\phi(y) \sim e^{-iky}(-(2ik)^3 + c2ik)e^{2iky} = \lambda e^{iky} .$$

It follows that $b = 0$ and $a = \lambda$, from which (B.5) follows. □

Differentiation of (B.5) yields the following for $Re\lambda > 0$. (For general λ, use analytic continuation.)

Corollary B.2. *Let $\lambda \in \Omega_0 = \mathbb{C}\backslash(-\infty, \lambda_*]$ (see Sect. 2), and let $\mu_1(\lambda)$ denote the root of $\mu^3 - c\mu + \lambda = 0$ with smallest real part. Then,*

$$Y^+(y, \lambda) = \mu_1(\lambda)^{-1}\partial_y[f(y, \mu_1(\lambda)/2i)^2] .$$

At this stage we can complete the proof of Theorem 3.1 (a), by determining $D(\lambda)$ explicitly for KdV. For KdV we have $z(y) = 3c\,\mathrm{sech}^2(\frac{1}{2}\sqrt{c}y)$. Corollary B.2 implies that we need to solve the scattering problem

$$-\partial_y^2 f - \frac{1}{2}c\,\mathrm{sech}^2(\frac{1}{2}\sqrt{c}y)f = k^2 f,\ e^{-iky}f(y, k) \to 1 \text{ as } y \to +\infty , \tag{B.6}$$

for $k = \mu_1(\lambda)/2i$. The solution of (B.6) is well-known to be

$$f(y, k) = Ae^{iky}\left[\frac{2ik}{\sqrt{c}} - \tanh(\frac{1}{2}\sqrt{c}y)\right] , \tag{B.7}$$

where the constant $A = \left(\frac{2ik}{\sqrt{c}} - 1\right)^{-1}$ has been chosen to satisfy the asymptotic condition in (B.3). (For a derivation of this result see the remark concluding this appendix.) Part (a) of Theorem 3.1 now follows from examining the expression for $Y^+(y, \lambda)$ given in Corollary B.2, in the limit $y \to -\infty$, namely

$$Y^+(y, \lambda) = \frac{1}{\mu_1}\partial_y\left[e^{\mu_1 y}\left(\frac{\mu_1 - \sqrt{c}\tanh(\frac{1}{2}\sqrt{c}y)}{\mu_1 - \sqrt{c}}\right)^2\right] . \tag{B.8}$$

The reduction to a second order Schrödinger scattering problem for mKdV is more involved and requires the complex Miura transform, as originally given by Miura [M].

Lemma B.3. *Let $\rho(x, t)$ be a solution of* mKdV. *Then $\psi(x, t)$ is a solution of* KdV, *where*

$$\psi = -i\sqrt{6}\partial_x\rho + \rho^2 . \tag{B.9}$$

Proof. By explicit calculation, $\psi_t + \psi\psi_x + \psi_{xxx} = (-i\sqrt{6}\,\partial_x + 2\rho)(\rho_t + \rho^2\rho_x + \rho_{xxx})$. □

If, in particular, we choose $\rho(x, t)$ to be a solitary wave solution of (mKdV),

$$\rho_0(x, t) = \sqrt{6c}\,\operatorname{sech}\sqrt{c}(x - ct)\,, \tag{B.10}$$

then we have:

Corollary B.4. *The* KdV *equation admits the complex solitary wave solution*

$$\psi_0(y) = 6c(\operatorname{sech}^2\sqrt{c}y + i\operatorname{sech}\sqrt{c}y\tanh\sqrt{c}y)\,, \tag{B.11}$$

where $y = x - ct$, for any $c > 0$.

To calculate $D(\lambda)$ for mKdV, we argue as follows, to reduce the problem to one for the complex KdV equation. Let $Re\lambda > 0$, and let $Y^+(y, \lambda)$ be the solution of

$$\partial_y(-\partial_y^2 + c - \rho_0^2)Y^+ = \lambda Y^+\,, \tag{B.12}$$

satisfying $Y^+(y, \lambda) \sim e^{\mu_1 y}$ as $y \to +\infty$. Recall that then $D(\lambda) = \lim_{y\to-\infty} Y^+(y, \lambda)\, e^{-\mu_1 y}$. We relate Y^+ to a solution W^+ of the linearized complex KdV equation via the linearized Miura transformation, as follows.

Lemma B.5. *Let $W^+(y, \lambda)$ be defined by*

$$-i\sqrt{6}\mu_1(\lambda)\cdot W^+(y, \lambda) = -i\sqrt{6}\,\partial_y Y^+(y, \lambda) + 2\rho_0(y)Y^+(y, \lambda)\,. \tag{B.13}$$

Then

$$\partial_y(-\partial_y^2 + c - \psi_0)W^+ = \lambda W^+\,, \tag{B.14}$$

and we have

$$W^+(y, \lambda) \sim e^{\mu_1 y} \text{ as } y \to +\infty, \quad W^+(y, \lambda) \sim D(\lambda)e^{\mu_1 y} \text{ as } y \to -\infty\,. \tag{B.15}$$

Proof. By explicit computation, with $\alpha = -i\sqrt{6}$ we find

$$(\partial_y^3 - \partial_y(c - \psi_0) + \lambda)\alpha W^+ = (\alpha\partial_y + 2\rho_0)[\partial_y^3 - \partial_y(c - \rho_0^2) + \lambda]Y^+\,.$$

The behavior in (B.15) follows since $\partial_y Y^+ \sim \mu_1 e^{\mu_1 y}$ (resp. $D(\lambda)\mu_1 e^{\mu_1 y}$) as $y \to +\infty$ (resp. $-\infty$), cf. [PW2]. □

The significance of Lemma B.5 is that Evans' function $D(\lambda)$ for the linearized mKdV equation (B.12) is the *same* as the corresponding Evans function for the linearized complex KdV equation (B.14). To determine the latter, we will find the Jost solutions of the associated Schrödinger equation with complex potential, and apply Lemma B.1 and Corollary B.2 to identify $W^+(y, \lambda)$.

The associated Schrödinger equation we must study is

$$-\partial_y^2 f - c(\operatorname{sech}^2\sqrt{c}y + i\operatorname{sech}\sqrt{c}y\tanh\sqrt{c}y)f = k^2 f\,. \tag{B.16}$$

The Jost solutions of this equation can be *explicitly* computed using a trick of Darboux (1882). Darboux observed that if $y(x)$ is the general solution of

$$y'' + (k^2 - u(x))y = 0\,, \tag{B.17}$$

and w denotes a particular solution with $k^2 = \beta$, satisfying

$$w'' + (\beta - u(x))w = 0\,, \tag{B.18}$$

then the general solution of an equation with different potential, namely

$$z'' + (k^2 - u(x) + 2(\ln w)'')z = 0\,, \tag{B.19}$$

is given by

$$z = y' - yw'/w\,. \tag{B.20}$$

To solve (B.16), take $u(x) = 0$ in (B.17)–(B.19), and observe that if $w(x) = e^x + ie^{-x}$, then $(\ln w)'' = 2(\operatorname{sech}^2 2x + i\operatorname{sech} 2x \tanh 2x)$. With this choice, (B.19) becomes

$$z'' + [k^2 + 4(\operatorname{sech}^2 2x + i\operatorname{sech} 2x\tanh 2x)]z = 0\,. \tag{B.21}$$

By now taking $y(x) = (ik-1)^{-1}e^{ikx}$ in (B.20), we obtain a solution of (B.21) given by

$$z(x) = (ik-1)^{-1}e^{ikx}[ik - \tanh 2x - i\operatorname{sech} 2x]\,. \tag{B.22}$$

This is easily related to the Jost solution of (B.16), via the change of variables $x = \sqrt{c}y/2$. This yields

$$f(y,k) = \left(\frac{2ik}{\sqrt{c}} - 1\right)^{-1} e^{iky}\left[\frac{2ik}{\sqrt{c}} - \tanh\sqrt{c}y - i\operatorname{sech}\sqrt{c}y\right]\,. \tag{B.23}$$

Now from Corollary B.2, we infer that the function W^+ in Lemma B.5 is given by

$$W^+(y,\lambda) = \frac{1}{\mu_1}\partial_y\left[e^{\mu_1 y}\left(\frac{\mu_1 - \sqrt{c}(\tanh\sqrt{c}y + i\operatorname{sech}\sqrt{c}y)}{\mu_1 - \sqrt{c}}\right)^2\right]\,. \tag{B.24}$$

By examining this expression in the limit $y \to -\infty$, we conclude that

$$D(\lambda) = \left(\frac{\mu_1 + \sqrt{c}}{\mu_1 - \sqrt{c}}\right)^2\,.$$

This completes the proof of Theorem 3.1. □

Remark. Darboux's trick can also be used to solve the scattering problem (B.6) which yielded, via Corollary B.2, $D(\lambda)$ for KdV. In this case we choose $u(x) = 0$ and $w(x) = \cosh(x)$. The expression (B.7) for the Jost solution now follows easily after scaling.

Acknowledgement. M.I.W. wishes to express his gratitude to the Institute for Advanced Study in Princeton, New Jersey and the Université de Paris-Sud, Orsay, for hospitality provided during the preparation of part of this paper.

References

[AS] Ablowitz, M.J., Segur, H.: Solitons and the Inverse Scattering Transform. SIAM, Philadelphia, 1981

[AGJ] Alexander, J., Gardner, R., Jones, C.K.R.T.: A topological invariant arising in the stability analysis of traveling waves. J. Reine Angew. Math. **410**, 167–212 (1990)

[Be] Benjamin, T.B.: The stability of solitary waves. Proc. Roy. Soc. Lond. **A328**, 153–183 (1972)

[Ber] Berryman, J.G.: Stability of solitary waves in shallow water. Phys. Fluids **19**, 771–777 (1976)

[Bo] Bona, J.L.: On the stability of solitary waves. Proc. Roy. Soc. Lond. **A344**, 363–374 (1975)

[BSc] Bona, J.L., Scott, R.: Solutions of the Korteweg–de Vries equation in fractional order Sobolev spaces. Duke Math J. **43**, 87–99 (1976)

[BSm] Bona, J.L., Smith, R.: The initial value problem for the Korteweg–de Vries equation. Phil. Trans. Roy. Soc. Lond. **A278**, 555–601 (1975)

[BSo] Bona, J.L., Soyeur, A.: On the stability of solitary wave solutions of model equations for long waves. J. Nonlin. Science

[BSS] Bona, J.L., Souganidis, P.E., Strauss, W.A.: Stability and instability of solitary waves. Proc. Roy. Soc. Lond. **A411**, 395–412 (1987)

[CL] Coddington, E., Levinson, N.: Theory of Ordinary Differential Equations. New York: McGraw-Hill, 1955

[C] Coppel, W.A.: Stability and asymptotic behavior of differential equations. Boston: D.C. Heath and Co., 1965

[CS] Crandall, M.G., Souganidis, P.E.: Convergence of difference approximations of quasilinear evolution equations. Nonl. Anal. TMA **10**, 425–445 (1986)

[CH1] Crawford, J.D., Hislop, P.D.: Application of the method of spectral deformation to the Vlasov Poisson system. Ann. Phys. **189**, 265–317 (1989)

[CH2] Crawford, J.D., Hislop, P.D.: Application of the method of spectral deformation to the Vlasov Poisson system II. Mathematical results. J. Math. Phys. **30**, 2819–2837 (1989)

[DJ] Drazin, P.G., Johnson, R.S.: Solitons: An introduction. Cambridge: Cambridge University Press, 1989

[E] Evans, J.W.: Nerve axon equations, IV: The stable and unstable impulse. Indiana Univ. Math. J. **24**, 1169–1190 (1975)

[GGKM1] Gardner, C.S., Greene, J.M., Kruskal, M.D., Miura, R.M.: Method for solving the Korteweg–de Vries equation. Phys. Rev. Lett. **19**, 1095–1097 (1967)

[GGKM2] Gardner, C.S., Greene, J.M., Kruskal, M.D., Miura, R.M.: Korteweg–de Vries equation and generalizations. VI. Methods for exact solution. Commun. Pure Appl. Math. **27**, 97–133 (1974)

[GKZ] Gesztesy, F., Karwowski, K., Zhao, Z.: New types of soliton solutions. Bull. Am. Math. Soc. **27**, 266–272 (1992)

[GT] Ginibre, J., Tsutsumi, Y.: Uniqueness for the generalized Korteweg–de Vries equation. SIAM J. Math. Anal. **20**, 1388–1425 (1989)

[H] Henry, D.: Geometric Theory of Semilinear Parabolic Equations. Lect. Notes in Math. **840**, New York: Springer, 1981

[JK] Jeffrey, A., Kakutani, T.: Weakly nonlinear dispersive waves: A discussion centered about the Korteweg–de Vries equation. SIAM Review **14**, 582–643 (1972)

[K1] Kato, T.: Quasilinear equations of evolution, with applications to partial differential equations. In: Spectral Theory and Differential Equations, A. Dold and B. Eckmann, (eds.), Lect. Notes in Math. **448**, New York: Springer, 1975

[K2] Kato, T.: Perturbation Theory for Linear Operators. 2nd ed., New York: Springer, 1976

[K3] Kato, T.: On the Cauchy problem for the (generalized) Korteweg–de Vries equation. Stud. in Appl. Math. Suppl. Ser. **8**, New York: Academic Press, 1983

[Ke] Keener, J.P.: Principles of Applied Mathematics. Redwood City: Addison-Wesley, 1988

[KM] Keener, J.P., McLaughlin, D.W.: Solitons under perturbations. Phys. Rev. **A 16**, 777–790 (1977)

[KA] Kodama, Y., Ablowitz, M.J.: Perturbations of solitons and solitary waves. Stud. in Appl. Math. **64**, 225–245 (1981)

[KPV] Kenig, C., Ponce, G., Vega, L.: Well-posedness of the initial value problem for the Korteweg–de Vries equation. J.A.M.S. **4**, 323–347 (1991)

[L] Lax, P.D.: Integrals of nonlinear equations of evolution and solitary waves. Commun. Pure Appl. Math. **21**, 467–490 (1968)

[LP] Lax, P.D., Phillips, R.S.: Scattering Theory. New York: Academic Press, 1967

[LS] Laedke, E.W., Spatschek, K.H.: Stability theorem for KdV type equations. J. Plasma Phys. **32**, 263–272 (1984)

[M] Miura, R.M.: Korteweg–de Vries equation and generalizations I. A remarkable explicit nonlinear transformation. J. Math. Phys. **9**, 1202–1204 (1968)

[Ne] Newell, A.C.: Near-integrable systems, nonlinear tunneling and solitons in slowly changing media. In: Nonlinear Evolution Equations Solvable by the Inverse Spectral Transform. F. Calogero, (ed.), London: Pitman, 1978, pp. 127–179

[Ni] Nirenberg, L.: Topics in Nonlinear Functional Analysis. Courant Institute Lecture Notes, 1974

[P] Pazy, A.: Semigroups of Linear Operators and Applications to Partial Differential Equations. Appl. Math. Sci. **44**, New York: Springer, 1983

[PW1] Pego, R.L., Weinstein, M.I.: On asymptotic stability of solitary waves. Phys. Lett. A **162**, 263–268 (1992)

[PW2] Pego, R.L., Weinstein, M.I.: Eigenvalues, and instabilities of solitary waves. Phil. Trans. Roy. Soc. Lond. **A340**, 47–94 (1992)

[PW3] Pego, R.L., Weinstein, M.I.: Evans' function, Melnikov's integral and solitary wave instabilities. In: Differential Equations with Applications to Mathematical Physics. (eds.) W.F. Ames, J.V. Herod, E.M. Harrell, Orlando: Academic Press, 1993

[RS3] Reed, M., Simon, B.: Methods of Modern Mathematical Physics III: Scattering Theory. New York: Academic Press, 1979

[RS4] Reed, M., Simon, B.: Methods of Modern Mathematical Physics IV: Analysis of Operators. New York: Academic Press, 1978

[Sa] Sattinger, D.H.: On the stability of waves of nonlinear parabolic systems. Adv. in Math. **22**, 312–355 (1976)

[Sc] Schuur, P.C.: Asymptotic Analysis of Soliton Problems. Lect. Notes in Math. **1232**, Berlin New York: Springer 1986

[SW1] Soffer, A., Weinstein M.I.: Multichannel nonlinear scattering theory for nonintegrable equations. In: Integrable Systems and Applications. Balabane, M., Lochak, P. and Sulem, C. (eds.), Springer Lect. Notes in Physics, **342**, 1988

[SW2] Soffer, A., Weinstein, M.I.: Multichannel nonlinear scattering theory for nonintegrable equations. Commun. Math. Phys. **133**, 119–146 (1990)

[SW3] Soffer, A., Weinstein M.I.: Multichannel nonlinear scattering theory for nonintegrable equations II: The case of anisotropic potentials and data. J. Diff. Eq. **98**, 376–390 (1992)

[T] Titchmarsh, E.C.: Eigenfunction Expansions Associated with Second-Order Differential Equations. Oxford: Oxford University Press, 1946

[V] Vainberg, B.: Asymptotic Methods in Equations of Mathematical Physics. New York: Gordon and Breach, 1989

[W1] Weinstein, M.I.: On the solitary wave of the generalized Korteweg–de Vries equation. In: Proc. Santa Fe Conference on Nonlinear PDE, July 1984, Lectures in Appl. Math. **23**, 1986

[W2] Weinstein, M.I.: Modulational stability of ground states of nonlinear Schrödinger equations. SIAM J. Math. Anal. **16**, 472–491 (1985)

[W3] Weinstein, M.I.: Lyapunov stability of ground states of nonlinear dispersive evolution equations. Commun. Pure Appl. Math. **39**, 51–68 (1986)

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

Communicated by T. Spencer