Proc. R. Soc. Lond. A **411**, 395–412 (1987)
Printed in Great Britain

Stability and instability of solitary waves of Korteweg–de Vries type

By J. L. Bona[1], P. E. Souganidis[2] and W. A. Strauss[2]

[1] *Department of Mathematics, The Pennsylvania State University, University Park, Pennsylvania* 16802, *U.S.A.*

[2] *Lefschetz Center for Dynamical Systems, Division of Applied Mathematics, Brown University, Providence, Rhode Island* 02912, *U.S.A.*

(*Communicated by T. B. Benjamin, F.R.S. – Received* 8 *October* 1987)

Considered herein are the stability and instability properties of solitary-wave solutions of a general class of equations that arise as mathematical models for the unidirectional propagation of weakly nonlinear, dispersive long waves. Special cases for which our analysis is decisive include equations of the Korteweg–de Vries and Benjamin–Ono type. Necessary and sufficient conditions are formulated in terms of the linearized dispersion relation and the nonlinearity for the solitary waves to be stable.

1. Introduction

This paper is concerned with the stability and instability properties of solitary-wave solutions $u = \phi(x-ct)$ of a general class of evolution equations of the form

$$u_t + u_x - Mu_x + f(u)_x = 0, \tag{1.1}$$

where subscripts denote partial differentiation, $u = u(x,t)$ is a real-valued function of the two real variables x and t, M is a constant coefficient pseudodifferential operator of order $\mu \geqslant 1$, and f is a general function. Such equations are mathematical models for the unidirectional propagation of weakly nonlinear, dispersive, long waves. In this application, u is an amplitude or a velocity, x is proportional to distance in the direction of propagation and t is proportional to elapsed time. The question raised here concerns specifying sharp conditions on M and f for which the solitary wave is stable.

A prototypical example of (1.1) is the generalized K.d.V. equation

$$u_t + u_x + u_{xxx} + (u^p)_x = 0, \tag{1.2}$$

wherein $\mu = 2$. If $p = 2$, this is the equation for surface water waves in a canal, derived by Korteweg & de Vries (1895). For $p = 2$ or 3, (1.2) is amenable to solution by the famous inverse-scattering theory of Gardner *et al.* (1967), and it may thereby be inferred that the solitary-wave solutions called solitons are stable (cf. Scharf & Wreszenski (1981) or Eckhaus & Schuur (1983) for recent accounts in this direction). For $p > 3$, the work of McLeod & Olver (1983) shows that (1.1) is not exactly solvable by an inverse-scattering transform, and numerical evidence (see Fornberg & Whitham 1978; Bona *et al.* 1986) confirms that the solitary waves are no longer solitons either. It is known, however, that the solitary-wave solution of (1.2) is stable if $1 < p < 5$ (see Benjamin 1972; Bona 1975; Weinstein 1987 and Albert *et al.* 1987). On the other hand, numerical simulations of solutions of (1.2) indicate that its solitary-wave solutions are unstable if $p \geqslant 5$, and in fact, that neighbouring solutions emanating from smooth initial data appear to form singularities in finite time. Although the blowing up of solutions is not demonstrated here, the first proof of instability is provided.

Another important model in the theory of long waves occurs when $\mu = 1$, which leads to the so-called Benjamin–Ono equation

$$u_t + u_x + Hu_{xx} + (u^p)_x = 0, \tag{1.3}$$

where H is the Hilbert transform. The solitary-wave solutions of (1.3) are proved to be stable if $1 < p < 3$ and unstable if $p \geqslant 3$. (The stability for $1 < p < 3$ was already established in Bennett *et al.* (1983) and Albert *et al.* (1987).)

In general, certain natural conditions on M, f, and ϕ are imposed, and under these conditions the solitary-wave solutions of (1.1) are proved to be stable or unstable depending upon the speed c of the wave. The principal result of the forthcoming analysis may be roughly understood as follows. For smooth functions g that vanish suitably at $\pm\infty$, define

$$\left.\begin{aligned} E(g) &= \int_{-\infty}^{\infty} \{\tfrac{1}{2}gMg - \tfrac{1}{2}g^2 - F(g)\}\,\mathrm{d}x, \\ V(g) &= \frac{1}{2}\int_{-\infty}^{\infty} g^2\,\mathrm{d}x, \\ I(g) &= \int_{-\infty}^{\infty} g\,\mathrm{d}x, \end{aligned}\right\} \tag{1.4}$$

where $F' = f$ and $F(0) = 0$. If $u = u(x,t)$ is an appropriately smooth solution of (1.1) then $E(u)$, $V(u)$, and $I(u)$ are independent of the temporal variable t; that is, E, V, and I are invariants of the motion generated by equation (1.1). Denote the solitary-wave solution ϕ of speed c by ϕ_c to emphasize its dependence on the parameter c. The main result then states that ϕ_c is stable *if and only if*

$$d(c) = E(\phi_c) + cV(\phi_c)$$

is a convex function of c. This condition is equivalent to the sufficient condition given by Shatah (1983) and by Weinstein (1987) in their analyses of the stability of ground states for certain nonlinear evolution equations. In case $f(u) = u^p$ is a pure power nonlinearity, $d(c)$ is convex if and only if $p < 2\mu + 1$.

The methods of analysis used herein derive mainly from those pioneered by Shatah & Strauss (1985) and Grillakis *et al.* (1987) in their study of bound-state solutions of nonlinear evolution equations. The abstract result of Grillakis *et al.* (1987) does not apply directly, however, because the range of the operator $\partial/\partial x$ which appears prominently in (1.1) does not include all smooth functions v, but only those for which $I(v) = 0$, i.e. functions that may be interpreted to have zero total mass if the dependent variable in the problem is proportional to the amplitude of a physical disturbance. One of the main new ideas in the approach taken here is to make serious use of the 'mass' invariant $I(u)$ defined in (1.4). The invariance of $I(u)$ insures the convergence of a certain improper integral which is used as a Lyapunov functional in our proof. Another related step is an estimation of integrals of the form $\int_{-\infty}^{a} u\,dx$.

The paper is organized as follows. A discussion of the evolution equation (1.1) and its three invariant functionals (1.4) is given in §2, along with a proof of the mass-related results to which reference was just made. Section 3 is devoted to explication of the solitary-wave solutions of (1.1). In particular, they are regarded as critical points of E subject to the constraint that V be constant. In the unstable case these critical points are not local minima, but rather saddle points, whereas in the stable case they comprise true local minima, as is remarked in §5. The analysis is especially simple when the nonlinearity is a pure power, but the general case is handled conclusively as well. In §4 the aforementioned Lyapunov functional is constructed and instability is proved following the main lines laid out by Shatah & Strauss (1985) and Grillakis *et al.* (1987). Finally, in §5 stability is established in the convex case. This result is essentially similar to several to which reference has already been made, but the form of nonlinearity encompassed by our theory is more general than was considered previously.

2. The evolution equation (1.1)

Because the stability or instability in view here refers to perturbations of the solitary-wave profile itself, a study of the initial-value problem for (1.1) is indicated. That is, we ask for a solution u of (1.1) defined for all x and $t \geqslant 0$ such that $\lim_{t\to 0} u(x,t) = u_0(x)$ where u_0 is a specified function. This corresponds to specifying the entire wave profile at some given instant of time and then inquiring as to the further evolution of the disturbance by application of the equation of motion.

Let $\widehat{Mu}(\xi) = |\xi|^{\mu}\,\hat{u}(\xi)$ with fixed $\mu \geqslant 1$ where the circumflexes denote Fourier transforms. (More general pseudodifferential operators could also be treated.) Let f be an odd, continuously differentiable, real valued function with $f(0) = f'(0) = 0$. In case $\mu = 1$, it is assumed that $|f(s)| = O(|s|^p)$ as $s \to \infty$ for some $p < \infty$, but for $\mu > 1$ this specification is not needed. The following theorem states that the initial-value problem for (1.1) is well posed in Hadamard's classical sense (see Abdelouhab *et al.* 1987).

Theorem 2.1. *Let $s > \frac{3}{2}$ and $f \in C^{s+1}$. For each $u_0 \in H^s(\mathbb{R})$, there exists $t_* = t_*(\|u_0\|_{H^s}) > 0$ and a unique solution $u \in C([0,t_*); H^s)$ of (1.1) with $u(\cdot,0) = u_0(\cdot)$. In addition, either $t_* = \infty$ or $\|u(\cdot,t)\|_{H^{\frac{1}{2}\mu}} \uparrow \infty$ as $t \uparrow t_*$.*

By a 'solution' of (1.1) we mean a solution in the weak sense. $H^s(\mathbb{R})$ is the Hilbert space of functions which together with their derivatives up to order s are square integrable. If $s > \mu + \frac{3}{2}$, then this weak solution is classical in the sense that each term in the equation is a continuous function and (1.1) is satisfied pointwise everywhere in the domain of the solution.

The next proposition establishes the temporal invariance of the functionals defined in (1.4).

Proposition 2.1. *Let s, f, and u_0 be as above. The unique solution of* (1.1) *with initial data u_0 satisfies $V(u(t)) = \text{const.}$ and $E(u(t)) = \text{const.}$ Moreover, if $s \geqslant 1+\mu$ and $\int_{-\infty}^{\infty} u_0(x)\,dx$ converges, then $I(u(t))$ converges for each t and is constant.*

Proof. The fact that V and E are constant follows from the construction in the existence theorem. Formally, they follow from multiplication of (1.1) by u and $Mu-f(u)-u$, respectively. If $s \geqslant \mu+1$ then $Mu \in C([0,t_*);H^1)$. Hence $Mu-f(u)-u$ is a continuous function of x and t and tends to zero as $|x| \to \infty$. Also, $\partial_t u = \partial_x(Mu-f(u)-u) \in C([0,t_*);L^2)$. Integrating (1.1) over the domain $\{(x,t): a \leqslant x \leqslant b, 0 \leqslant t \leqslant T\}$ yields

$$\int_0^T \int_a^b \partial_t u\,dx\,dt = \int_0^T \int_a^b \partial_x(Mu-u-f(u))\,dx\,dt.$$

As $a \to -\infty$, $b \to \infty$, the right-hand side tends to zero. Hence the improper integral $\int_{-\infty}^{\infty} u(x,T)\,dx$ exists and equals $\int_{-\infty}^{\infty} u_0(x)\,dx$ for each $T < t_*$. ∎

Having established that $I(u)$ is independent of t, interest is turned to estimating about how fast its tail near infinity grows with t. Such information will turn out to be essential in our proof of instability.

Theorem 2.2. *Let $u_0 \in H^{\mu+1}$ and $\theta = \int_{-\infty}^{\infty} (1+|x|)\,|u_0(x)|\,dx < \infty$. Assume that $f(s) = O(s^2)$ as $s \to 0$. If u is the solution of* (1.1) *corresponding to u_0, then*

$$\sup_{-\infty<z<\infty} \left| \int_z^\infty u(x,t)\,dx \right| \leqslant c(t^{\mu/(1+\mu)} + t^{-\mu/(1+\mu)}), \tag{2.1}$$

for $0 \leqslant t \leqslant t_1$, where c depends only on $\sup_{0 \leqslant t \leqslant t_1} \|u(t)\|_{H^{\frac{1}{2}\mu}} + \theta$, and t_1 is any time for which this expression is finite.

Proof. If $\gamma = \int_{-\infty}^{\infty} u_0(x)\,dx$, then $\gamma = \int_{-\infty}^{\infty} u(x,t)\,dx$ for $t \in [0,t_*)$ and therefore

$$\int_z^\infty u(x,t)\,dx = \gamma - \int_{-\infty}^z u(x,t)\,dx = \gamma - U(z,t), \tag{2.2}$$

where U is a solution of the initial-value problem

$$\left.\begin{aligned} U_t - MU_x + U_x + f(u) &= 0, \\ U(x,0) &= \int_{-\infty}^x u_0(y)\,dy. \end{aligned}\right\} \tag{2.3}$$

Let K be the fundamental solution of $Z_t - MZ_x = 0$; that is, K is the solution of

$$\left.\begin{aligned} K_t - MK_x &= 0, \\ K(x,0) &= \delta(x), \end{aligned}\right\} \tag{2.4}$$

where $\delta(x)$ is the Dirac delta function centred at 0. The scaling properties of (2.4) imply that K has the special form

$$K(x,t) = t^{-1/(1+\mu)}\,k(xt^{-1/(1+\mu)}) \quad \text{for} \quad t > 0, \tag{2.5}$$

where k is defined by

$$k(s) = (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{\infty} e^{i(s\xi - \xi|\xi|^\mu)}\,d\xi. \tag{2.6}$$

The solution U of (2.3) may be expressed in the form

$$U(x,t) = \int_{-\infty}^{\infty} K(x-y-t,t)\, U(y,0)\, \mathrm{d}y - \int_0^t \int_{-\infty}^{\infty} K(x-y-(t-\tau), t-\tau) \times f(u(y,\tau))\, \mathrm{d}y\, \mathrm{d}\tau. \quad (2.7)$$

To continue we need to comprehend certain detailed properties of K and k. These are stated in the following lemma, and proved after the theorem is established.

LEMMA 2.1. *The function k defined by* (2.6) *is bounded and $\int_{-\infty}^{\infty} k(s)\, \mathrm{d}s$ exists as an improper integral.*

Let $\mathscr{H}$ be the Heaviside function. The assumption on u_0 implies that $\int_{-\infty}^{\infty} |U(x,0) - \gamma\mathscr{H}(x)|\, \mathrm{d}x < \infty$. Moreover, lemma 2.1, combined with (2.5) gives that

$$|K(z,t)| \leqslant Ct^{-1/(1+\mu)} \quad \text{for all} \quad z \in \mathbb{R} \quad \text{and} \quad t > 0, \quad (2.8)$$

where C is a constant, and that

$$\int_{-\infty}^{\infty} K(z,t)\, \mathrm{d}z \quad (2.9)$$

exists as an improper integral and is independent of t, for all $t > 0$. Combining the above remarks leads to the estimate

$$\begin{aligned}\left|\int_{-\infty}^{\infty} K(x-y-t,t)\, U(y,0)\, \mathrm{d}y\right| &= \left|\int_{-\infty}^{\infty} K(x-y-t,t)\, [\gamma\mathscr{H}(y) + U(y,0) - \gamma\mathscr{H}(y)]\, \mathrm{d}y\right| \\ &\leqslant |\gamma| \left|\int_{-\infty}^{p} k(s)\, \mathrm{d}s\right| + C \| U(\cdot,0) - \gamma\mathscr{H}(\cdot)\|_{L^1} \|K(\cdot,t)\|_{L^\infty} \\ &\leqslant C(1 + t^{-1/(1+\mu)}),\end{aligned}$$

for some constant $C > 0$, where $p = (x-t)\, t^{-1/(1+\mu)}$.

To estimate the second term on the right-hand side of (2.7), the case where $\mu > 1$ is considered separately from the case $\mu = 1$. Let $t_1 > 0$ be such that $c = \sup_{t\in[0,t_1]} \|u(t)\|_{H^{\frac{1}{2}\mu}} < \infty$. If $\mu > 1$, there exists a constant C, which depends on c such that $\sup_{t\in[0,t_1]} \|u(t)\|_{L^\infty} \leqslant C$. Because $f(s) = g(s)\, s^2$ with $g(s)$ bounded for s bounded, it is implied that $|f(u)| \leqslant c'|u|^2$ at least for $0 \leqslant t \leqslant t_1$. Putting these considerations together yields

$$\begin{aligned}\left|\int_0^t \int_{-\infty}^{\infty} K(x-(t-\tau)-y, t-\tau) f(u(y,\tau))\, \mathrm{d}y\, \mathrm{d}\tau\right| &\leqslant C \int_0^t \|K(\cdot, t-\tau)\|_{L^\infty} \|u(\tau)\|_{L^2}^2\, \mathrm{d}\tau \\ &\leqslant C\|u(\cdot,\tau)\|_{L^2}^2 \int_0^t (\tau-\tau)^{-1/(1+\mu)}\, \mathrm{d}\tau \\ &\leqslant Ct^{\mu/(1+\mu)}.\end{aligned}$$

If $\mu = 1$, it is assumed that $|f(s)| = O(|s|^p)$ as $|s| \to \infty$ for some p with $1 \leqslant p < \infty$. Additionally, recall that $H^{\frac{1}{2}}(\mathbb{R}) \subset L^q(\mathbb{R})$ for every $q \in [2,\infty)$. Combining these gives $\int_{-\infty}^{\infty} |f(u)\,(y,t)|\, \mathrm{d}y \leqslant C$ where $C = C(c)$. Then

$$\left|\int_0^t \int_{-\infty}^{\infty} K(x-y-(t-\tau), t-\tau) f(u(y,\tau))\, \mathrm{d}y\, \mathrm{d}\tau\right| \leqslant Ct^{\frac{1}{2}}.$$

The theorem is thus established in all cases.

Proof of lemma 2.1. The method of stationary phase is used. In showing that k is bounded or improperly integrable, the hardest range of the dependent variable to handle is s large and positive, and it is to this case that attention focuses. Observe that after a change of variables, k may be expressed as

$$k(s) = (2\pi)^{-\frac{1}{2}} s^{1/\mu} \int_{-\infty}^{\infty} e^{ixh(\xi)} \, d\xi,$$

where $$x = s^{1+1/\mu} \quad \text{and} \quad h(\xi) = \xi - \xi|\xi|^{\mu}.$$

The function h has two critical points, namely $\alpha = \pm(\mu+1)^{-1/\mu}$. Let $[a, b]$ be an interval (with either a or b possibly infinite) that does not contain any of the critical points of h. Integrating by parts twice gives the formula

$$\int_a^b e^{ixh(\xi)} \, d\xi = \frac{e^{ixh(\xi)}}{ixh'(\xi)} + \frac{e^{ixh(\xi)}}{x^2} \frac{h''(\xi)}{(h'(\xi))^2}\bigg]_a^b - \frac{1}{x^2} \int_a^b e^{ixh(\xi)} \left(\frac{h''}{(h')^3}\right)' d\xi.$$

Each of these terms separately tends to zero as $s \to +\infty$ and the improper integral with respect to s obviously exists.

Next observe that near a critical point α one may write $h(\xi) = h(\alpha) + (\xi-\alpha)^2 f(\xi)$ where f is analytic and $f(\alpha) \neq 0$. Consider the interval $[\alpha, \alpha+\epsilon]$ where ϵ is a fixed small number. After the change of variables $\xi = \alpha + (t/x)^{\frac{1}{2}}$, there appears the relation

$$k_2(s) \equiv s^{1/\mu} \int_{\alpha}^{\alpha+\epsilon} e^{ixh(\xi)} \, d\xi = \tfrac{1}{2} s^{1/\mu} x^{-\frac{1}{2}} e^{ixh(\alpha)} g(x),$$

where $$g(x) = \int_0^{\epsilon} \exp\{itf(\alpha + (t/x)^{\frac{1}{2}})\} \, t^{-\frac{1}{2}} \, dt.$$

An integration by parts shows that $g(x)$ is bounded as $x \to +\infty$. Hence

$$k_2(s) = O(s^{1/\mu} x^{-\frac{1}{2}}) = O(s^{\mu-\frac{1}{2}-1/2\mu}),$$

as $s \to \infty$. This is bounded because $\mu \geqslant 1$. As for the integral of k_2 with respect to s,

$$\int_{-\infty}^{\infty} k_2(s) \, ds = c \int_{-\infty}^{\infty} e^{ixh(x)} g(x) \, x^{-\frac{1}{2}} \, dx.$$

From repeated integrations by parts, it is easily seen that this integral converges. ∎

3. The solitary-wave solution of (1.1)

Consider a smooth solution of the evolution equation (1.1) of the form $U(x,t) = \phi_c(x-ct)$ with $c > 1$. So long as c is fixed we write ϕ for ϕ_c. The function ϕ satisfies the equation

$$-c\,\partial_x \phi - \partial_x M\phi + \partial_x \phi + \partial_x f(\phi) = 0. \tag{3.1}$$

Assume that $\phi(x)$ and $M\phi(x)$ decay to zero as $|x| \to \infty$, so that (3.1) implies

$$M\phi + c\phi - \phi - f(\phi) = 0. \tag{3.2}$$

It is important for the arguments put forth in §§4 and 5 to characterize ϕ in terms of the functionals

$$V(u) = \frac{1}{2}\int_{-\infty}^{\infty} u^2\,\mathrm{d}x, \tag{3.3}$$

and

$$E(u) = \int_{-\infty}^{\infty} \{\tfrac{1}{2}uMu - \tfrac{1}{2}u^2 - F(u)\}\,\mathrm{d}x, \tag{3.4}$$

introduced earlier. For this task, a natural functional–analytical framework is now introduced. Let $X = H^{\frac{1}{2}\mu}(\mathbb{R})$, denote by $\langle\cdot,\cdot\rangle$ the L^2-inner product, and let $\langle\cdot,\cdot\rangle_*$ denote the usual pairing between X and its dual space X^*, where X^* is identified with $H^{-\frac{1}{2}\mu}(\mathbb{R})$. (Of course if $u\in X$ and $v \in L^2$, then $\langle u, v\rangle_* = \langle u, v\rangle$.) The immediate goal is to consider V and E as mappings of X into $\mathbb{R}$. Indeed, it is immediate that V is a C^∞-mapping of X into $\mathbb{R}$, that $V'(u) = u$, and that $V''(u) =$ identity. (More precisely, $V'(u)v = \langle u, v\rangle_* = \langle u, v\rangle$ and $V''(u)(v,w) = \langle v, w\rangle$, for all $v, w\in X$.) As for E, if $\mu > 1$, then $X \subset L^\infty(\mathbb{R})$. Because $F(0) = F'(0) = 0$, we have $F(s) = O(s^2)$ as $s\to 0$. Thus $|\int F(u)\,\mathrm{d}x| < \infty$ for every $u\in X$. If $\mu = 1$, then $X\subset L^{q+1}(\mathbb{R})$ for every q with $1 \leqslant q < \infty$. The growth condition imposed on f in this case implies that $|F(u)| \leqslant c(|u|^2 + |u|^{q+1})$ for some $q\in[1,\infty)$. Thus $|\int F(u)\,\mathrm{d}x| \leqslant c[\|u\|_X^2 + \|u\|_X^{q+1}]$. Finally, $\int_{-\infty}^{\infty} u\,Mu\,\mathrm{d}x = \int |\xi|^\mu\,|\hat{u}|^2\,\mathrm{d}\xi < \infty$. It is therefore easy to check that $E'(u) \in X^*$, that $E''(u)$ exists, and that

$$E'(u) = Mu - u - f(u) \quad \text{and} \quad E''(u) = M - 1 - f'(u). \tag{3.5}$$

In view of the above, (3.2) may be written as

$$E'(\phi) + cV'(\phi) = 0. \tag{3.6}$$

Moreover, the linearized operator $\mathscr{L}_c = \mathscr{L}$ of $E' + cV'$ around ϕ_c is

$$\mathscr{L}_c = \mathscr{L} = M + c - 1 - f'(\phi) = E''(\phi) + cV''(\phi). \tag{3.7}$$

It is immediate that $\mathscr{L}: X\to X^*$ and $\langle \mathscr{L}u, v\rangle = \langle u, \mathscr{L}v\rangle$ for all $u, v\in X$. One may also define $\overline{\mathscr{L}} = (1+M)^{-\frac{1}{2}}\mathscr{L}(1+M)^{-\frac{1}{2}}$ with $D(\overline{\mathscr{L}}) = H^{2\mu}(\mathbb{R})$. Then $\overline{\mathscr{L}}$ is an unbounded, self-adjoint operator on $L^2(\mathbb{R})$. Finally, note that (3.1) implies

$$\mathscr{L}(\partial_x \phi) = 0. \tag{3.8}$$

We continue by stating our assumptions on ϕ and $\mathscr{L}$. These are the following.

There is an interval (c_1, c_2) with $1 \leqslant c_1 < c_2 \leqslant \infty$ such that for every $c\in(c_1,c_2)$ there exists a solution ϕ_c of (3.2). The curve $c\mapsto\phi_c$ is C^1 with values in $H^{1+\frac{1}{2}\mu}(\mathbb{R})$. Moreover, $\phi_c(x) > 0$, $\phi_c \in H^{3+\frac{1}{2}\mu}(\mathbb{R})$ and $(1+|x|)^{\frac{1}{2}}\,\mathrm{d}\phi_c/\mathrm{d}x \in L^1(\mathbb{R})$. (3.9)

Also,

the operator $\mathscr{L}_c$ has a unique, negative, simple eigenvalue, with eigenfunction χ_c, the zero eigenvalue (with eigenfunction $\partial_x\phi$) is simple, and all the rest of the spectrum of $\mathscr{L}$ is positive and bounded away from zero. Moreover, the mapping $c\mapsto\chi_c$ is a continuous curve with values in $H^{1+\frac{1}{2}\mu}(\mathbb{R})$, $\chi_c(x) > 0$ and $(1+|x|^{\frac{1}{2}})\,\chi_c \in L^1(\mathbb{R})$. (3.10)

Before we continue, we remark that the solitary wave $\phi = \phi_c$ and $d\phi_c/dc$ are in H^∞, as soon as they are in $H^{\frac{1}{2}\mu}$. This follows immediately from the equation and a simple *ab initio* argument. Next observe that (3.6) characterizes ϕ as a critical point of E subject to the constraint $V(u) = V(\phi)$. It transpires that ϕ is either a local minimum or a saddle point, and that the stability or not of ϕ is determined by which possibility occurs. Define $d(c) = E(\phi_c)+cV(\phi_c)$. Then a simple aspect of the function d determines the type of the critical point ϕ, as is demonstrated in the next theorem and lemma 5.2.

THEOREM 3.1. *Let c be fixed. If $d''(c) < 0$, then there exists a curve $\omega \to \psi_\omega$ which passes through ϕ_c, lies on the surface $V(u) = V(\phi_c)$, and on which $E(u)$ has a strict local maximum at $u = \phi_c$.*

Two proofs of theorem 3.1 will be given. The first is direct and is based on dilations, but requires an additional hypothesis if $\mu > 1$. (This hypothesis is always satisfied if $f(u) = u^p$.) The second, which is more general and abstract, is taken from Grillakis *et al.* (1987).

First proof. In case $\mu > 1$, we make the additional assumption that $(d^{\mu/(\mu+1)})''(c) > 0$. (This is satisfied in the key examples, as will be shown at the end of this section.) Define the dilated curve $\psi_\omega(x) = \phi_\omega(x/\sigma(\omega))$, where $\sigma(\omega) = V(\phi_c)/V(\phi_\omega)$. Then $\psi_c(x) = \phi_c(x)$ and

$$V(\psi_\omega) = \frac{1}{2}\int \phi_\omega^2\left(\frac{x}{\sigma(\omega)}\right)dx = \frac{\sigma(\omega)}{2}\int \phi_\omega^2(y)\,dy = \sigma(\omega)\,V(\phi_\omega) = V(\phi_c).$$

However,

$$\left.\begin{aligned} E(\psi_\omega)+\omega V(\psi_\omega) &= \int\left[\tfrac{1}{2}\psi_\omega M\psi_\omega - \tfrac{1}{2}\psi_\omega^2 - F(\psi_\omega) + \frac{\omega}{2}\psi_\omega^2\right]dx,\\ &= \frac{\sigma^{1-\mu}}{2}\int \phi_\omega M\phi_\omega + \sigma\int\left[-\tfrac{1}{2}\phi_\omega^2 - F(\phi_\omega) + \frac{\omega}{2}\phi_\omega^2\right]dx. \end{aligned}\right\} \tag{3.11}$$

Differentiating (3.11) with respect to ω and evaluating at $\omega = c$ leads to the relation

$$\begin{aligned} 0 &= \left\langle E'(\phi_c)+cV'(\phi_c), \left.\frac{d\psi_\omega}{d\omega}\right|_{\omega=c}\right\rangle \\ &= \sigma'(c)\left\{(1-\mu)\tfrac{1}{2}\langle\phi_c, M\phi_c\rangle + \int\left[-\tfrac{1}{2}\phi_c^2 - F(\phi_c) + \frac{c}{2}\phi_c^2\right]dx\right\} \\ &\quad + \left\langle M\phi_c - \phi_c - f(\phi_c) + c\phi_c, \left.\frac{d\phi_\omega}{d\omega}\right|_{\omega=c}\right\rangle. \end{aligned}$$

Because of (3.2), the last formula simplifies to

$$\sigma'(c)\{-\tfrac{1}{2}\mu\langle\phi_c, M\phi_c\rangle + d(c)\} = 0. \tag{3.12}$$

Differentiating $d(\omega) = E(\phi_\omega)+\omega V(\phi_\omega)$ with respect to ω gives,

$$d'(\omega) = \langle E'(\phi_\omega)+\omega V'(\phi_\omega), d\phi_\omega/d\omega\rangle + V(\phi_\omega) = V(\phi_\omega). \tag{3.13}$$

The definition of $\sigma(\omega)$ yields

$$\sigma'(\omega) = -[V(\phi_c)/(d'(\omega))^2]\,d''(\omega).$$

Therefore, by assumption,

$$\sigma'(c) = -d''(c)/d'(c) > 0. \tag{3.14}$$

So (3.12) implies the 'dilation identity'

$$d(c) = \tfrac{1}{2}\mu\langle\phi_c, M\phi_c\rangle. \tag{3.15}$$

Returning to (3.11), we find $E(\psi_\omega)+\omega V(\psi_\omega)$ expressed in the simple form

$$E(\psi_\omega)+\omega V(\psi_\omega) = (\sigma^{1-\mu}-\sigma)\tfrac{1}{2}\langle\phi_\omega, M\phi_\omega\rangle+\sigma d(\omega),$$
$$= [(\sigma^{1-\mu}-\sigma)/\mu+\sigma]\,d(\omega).$$

The derivatives with respect to ω of the last identity lead to the formulae

$$(\mathrm{d}/\mathrm{d}\omega)\,E(\psi_\omega)+V(\phi_c) = [(\mu-1)/\mu][1-\sigma^{-\mu}]\,\sigma' d+(1/\mu)[\sigma^{1-\mu}+(\mu-1)\,\sigma]\,d'$$

and

$$\mathrm{d}^2/\mathrm{d}\omega^2\,E(\psi_\omega) = (\mu-1)\,(\sigma')^2\,d+d'' = d''[(\mu-1)\,dd''/(d')^2+1]$$
$$= [(\mu-1)^2/\mu]\,d^{(\mu-2)/(\mu-1)}\,(d')^{-2}\,d''\,(d^{\mu/(\mu-1)})''$$

except when $\mu = 1$ where

$$(\mathrm{d}/\mathrm{d}\omega)\,E(\psi_\omega)+V(\psi_\omega) = d' \quad\text{and}\quad (\mathrm{d}^2/\mathrm{d}\omega^2)\,E(\psi_\omega) = d''.$$

In any case, if these formulae are evaluated at $\omega = c$, it is found that

$$(\mathrm{d}/\mathrm{d}\omega)\,E(\psi_\omega)|_{\omega=c} = 0 \quad\text{and}\quad (\mathrm{d}^2/\mathrm{d}\omega^2)\,E(\psi_\omega) < 0,$$

because $d(c) > 0$ by (3.15), and by the assumption that $d''(c) < 0$. (If $\mu > 1$, the additional assumption mentioned at the outset comes into play.) ∎

Second proof. Let χ_c be the unique, negative eigenfunction of $\mathscr{L}_c$ whose existence was postulated in (3.10). For ω near c, define $\psi_\omega = \varphi_\omega+s(\omega)\,\chi_c$ with $s(\omega)$ determined by the requirements $s(c) = 0$, $\psi_c = \phi_c$ and $V(\psi_\omega) = V(\phi_c)$. That such a function $s(\omega)$ exists for ω near c follows from the implicit function theorem, once it is remarked that

$$\frac{\partial}{\partial s}V(\phi_\omega+s\chi_c)|_{\substack{s=0\\ \omega=c}} = \int\phi_c(x)\,\chi_c(x)\,\mathrm{d}x,$$

and the latter integral is strictly positive because both ϕ_c and χ_c are. A straightforward calculation gives

$$\frac{\mathrm{d}^2}{\mathrm{d}\omega^2}E(\psi_\omega) = \frac{\mathrm{d}^2}{\mathrm{d}\omega^2}[E(\psi_\omega)+\omega V(\psi_\omega)] = \left\langle E'(\psi_\omega)+\omega V'(\psi_\omega), \frac{\mathrm{d}^2\psi_\omega}{\mathrm{d}\omega^2}\right\rangle$$
$$+\left\langle E''(\psi_\omega)+\omega V''(\psi_\omega)]\frac{\mathrm{d}\psi_\omega}{\mathrm{d}\omega}, \frac{\mathrm{d}\psi_\omega}{\mathrm{d}\omega}\right\rangle.$$

At $\omega = c$ this simplifies to

$$(\mathrm{d}^2/\mathrm{d}\omega^2)\,E(\psi_\omega)|_{\omega=c} = \langle\mathscr{L}_c y, y\rangle,$$

where

$$y = \frac{\mathrm{d}\psi_\omega}{\mathrm{d}\omega}\bigg|_{\omega=c} = \frac{\mathrm{d}\phi_c}{\mathrm{d}c}+s'(c)\,\chi_c. \tag{3.16}$$

The constancy of V along ψ_ω means that

$$0 = \frac{\mathrm{d}}{\mathrm{d}\omega} V(\psi_\omega)|_{\omega=c} = \int y\phi_c \,\mathrm{d}x = \left\langle \frac{\mathrm{d}\varphi_c}{\mathrm{d}c}, \phi_c \right\rangle + s'(c) \langle \chi_c, \phi_c \rangle$$

and therefore

$$\int y\phi_c \,\mathrm{d}x = 0 \quad \text{and} \quad d''(c) = -s'(c) \langle \chi_c, \varphi_c \rangle. \tag{3.17}$$

Also

$$\mathscr{L}_c y = \mathscr{L} \,\mathrm{d}\phi_c/\mathrm{d}c + s'(c) \mathscr{L}_c \chi_c = -\phi_c + s'(c) \mathscr{L}_c \chi_c,$$

so by (3.16) and (3.17)

$$\begin{aligned} \langle \mathscr{L}_c y, y \rangle &= s'(c) \langle \mathscr{L}_c \chi_c, y \rangle = s'(c) \langle \chi_c, \mathscr{L}_c y \rangle \\ &= -s'(c) \langle \chi_c, \phi \rangle + [s'(c)]^2 \langle \chi_c, \mathscr{L}_c \chi_c \rangle \\ &= d''(c) + [s'(c)]^2 \langle \chi_c, \mathscr{L}_c \chi_c \rangle \end{aligned}$$

and the latter expression is negative because $d''(c) < 0$ by assumption and $\langle \chi_c, \mathscr{L}_c \chi_c \rangle < 0$ from (3.10).

This completes the proof of theorem 3.1. ■

We continue by calculating $d(c)$ explicitly in the case where $f(u) = u^p$. This will facilitate the determination of the exact range of p for which $d''(c) < 0$. Moreover, it will appear that the extra assumption made for the first proof of theorem 3.1 is always valid for such a pure power nonlinearity.

For every $c > 1$ the solution ϕ_c of (3.2) can be written as

$$\phi_c(x) = (c-1)^{1/(p-1)} v((c-1)^{1/\mu} x),$$

where v is the solution of

$$Mv + (c-1)v - v^p = 0.$$

Because of this, (3.15) allows us to infer that

$$d(c) = \tfrac{1}{2}\mu(c-1)^{2/(p-1)+1-1/\mu} \langle v, Mv \rangle, \tag{3.18}$$

from which we immediately deduce the following corollary.

Corollary. *For every $c > 1$, the solitary-wave solution of* (1.1) *with pure power nonlinearity has the property that*

(i) $d''(c) < 0$ *if and only if* $p < 2\mu + 1$, *and*

(ii) $(d^{\mu/(\mu-1)})''(c) > 0$.

4. Instability

The theory relating to instability of solitary-wave solutions of equation (1.1) is developed in this section. We begin by specifying the precise way in which stability, and its negation, instability, is to be interpreted. For any real number s let τ_s denote translation by s acting on functions of one real variable. That is, $(\tau_s f)(x) = f(x+s)$ for all $x \in \mathbb{R}$. For $\epsilon > 0$, consider the 'tube'

$$U_\epsilon = \{u \in X : \inf_s \|u - \tau_s \phi_c\|_X < \epsilon\}. \tag{4.1}$$

The set U_ϵ is a neighbourhood of the collection of all translates of ϕ_c.

Definition 4.1. The solitary wave ϕ_c is *stable* if and only if for every $\epsilon > 0$ there exists $\delta > 0$ such that if $u_0 \in U_\delta$, then $u(\cdot, t) \in U_\epsilon$ for all $t \in \mathbb{R}$. The solitary wave ϕ_c is *unstable* if ϕ_c is not stable.

Because the set of all translates of ϕ_c comprises its orbit under the action of the evolution equation (1.1), the stability specified here corresponds to what is often called orbital stability. It states that if a solution of (1.1) is globally defined and if it resembles sufficiently closely a solitary wave at one instant of time, then it always was and always will be nearly a solitary wave, at least so far as its shape measured in the norm on X is concerned.

Here is the principal result of the paper.

THEOREM 4.1. *The solitary wave ϕ_c is stable if and only if $d''(c) > 0$.*

Throughout this section we assume a speed c larger than one is fixed and ϕ_c is abbreviated to just ϕ. Neighbourhoods U_ϵ that appear will refer to the fixed solitary-wave solution ϕ. The following technical lemma will be needed in the proof of theorem 4.1.

LEMMA 4.1. *There exist $\epsilon > 0$ and a unique C^1 map $\alpha: U_\epsilon \to \mathbb{R}$ such that for every $u \in U_\epsilon$ and $r \in \mathbb{R}$,*

(i) $\langle u(\cdot + \alpha(u)), \partial_x \phi \rangle = 0$,

(ii) $\alpha(u(\cdot + r)) = \alpha(u) - r$,

(iii) $\alpha'(u) = \dfrac{\partial_x \phi(\cdot - \alpha(u))}{\int_{-\infty}^{\infty} u(x)\, \partial_x^2 \phi(x - \alpha(u))\, dx}$.

Proof. Consider the functional

$$(u, \alpha) \to \int u(x + \alpha)\, \partial_x \phi(x)\, dx,$$

defined on pairs $u \in L^2(\mathbb{R})$ and $\alpha \in \mathbb{R}$. Its derivative with respect to α at $\alpha = 0$ and $\mu = \phi$ is $\int (\partial_x \phi)^2\, dx$, which is non-zero. By the implicit function theorem, there is a unique C^1 functional $\alpha(u)$ satisfying (i) in a neighbourhood of ϕ. By translation invariance, $\alpha(u)$ can be uniquely extended to a tube of the form U_ϵ for $\epsilon > 0$ small enough. By (i), $u(\cdot + \alpha(u)) = u(\cdot + r + (\alpha(u) - r))$ is orthogonal to $\partial_x \phi$. Therefore by the uniqueness of $\alpha(u)$, guaranteed by the conclusion of the implicit function theorem, $\alpha(u) - r = \alpha(u(\cdot + r))$ so that (ii) holds. Finally, we calculate $\alpha'(u)$. A change of variables converts (i) to the relation

$$\int_{-\infty}^{\infty} u(x)\, \partial_x \varphi(x - \alpha(u))\, dx = 0.$$

Differentiating with respect to $u \in X$ gives

$$\partial_x \varphi(\cdot - \alpha(u)) - \int_{-\infty}^{\infty} u(x)\, \partial_x^2 \phi(x - \alpha(u))\, dx\, \alpha'(u) = 0.$$

This leads to (iii) because $u \in H^{\frac{1}{2}\mu}$ and $\partial_x^2 \phi \in H^{1+\frac{1}{2}\mu}$. The proof is thus complete. Notice that at $u = \phi$,

$$\langle u, \partial_x^2 \phi \rangle = \langle \phi, \partial_x^2 \phi \rangle = -\langle \partial_x \phi, \partial_x \phi \rangle < 0.$$

Also $\langle \partial_x \alpha'(u), u \rangle = 1$ at $u = \phi$, and α is a C^2 functional on X because $\partial_x^3 \phi \in X$. ■

The development is continued by defining an auxiliary operator B which will play a critical role in the proof of instability. Let

$$y = d\psi_\omega/d\omega|_{\omega=c},$$

where ψ_ω is either one of the curves defined in §3. It follows from the calculations given in the proof of theorem 3.1 that

$$\langle \mathscr{L}_c y, y\rangle < 0 \quad \text{and} \quad \langle y, \phi\rangle = 0.$$

Definition 4.1. For $u \in U_\epsilon$, define $B(u)$ by the formula

$$B(u) = y(\cdot - \alpha(u)) - \langle u, y(\cdot - \alpha(u))\rangle\, \partial_x \alpha'(u).$$

By lemma 3.1, B may also be expressed as

$$B(u) = y(\cdot - \alpha(u)) - \frac{\langle u, y(\cdot - \alpha(u))\rangle}{\langle u, \partial_x^2 \phi(\cdot - \alpha(u))\rangle} \partial_x^2 \phi(\cdot - \alpha(u)). \tag{4.2}$$

The next proposition summarizes some important properties of B.

PROPOSITION 4.1. *B is a C^1 function from U_ϵ into X. Moreover, B commutes with translations, $B(\phi) = y$ and $\langle B(u), u\rangle = 0, \forall u \in U_\epsilon$.*

Proof. It is clear that $B(u) \in X$ because $y \in X$ and $\partial_x^2 \phi \in H^{1+\frac{1}{2}\mu}$. Also $B(u)\, w \in X$ for every $u \in U_\epsilon$ and $w \in X$ because $\partial_x^3 \phi \in X$ and $\partial_x y \in X$. The latter follows from the fact that $\partial_x(d\phi_c/dc)$ and $\partial_x \chi$ both lie in X because, by assumption, $d\phi/dc$ and χ are members of $H^{1+\frac{1}{2}\mu}$. The fact that B is a C^1 function follows from a straightforward, explicit calculation, which is left for the reader's amusement. The next task is to calculate $B(\phi)$. Because $\alpha(\phi) = 0$ and $\langle \phi, y\rangle = 0$, it is easily determined that

$$B(\phi) = y + \frac{\langle \phi, y\rangle}{\langle \partial_x \phi, \partial_x \phi\rangle} \partial_x^2 \phi = y.$$

Also, (4.2) obviously implies that

$$\langle B(u), u\rangle = 0.$$

Finally,

$$\begin{aligned} B(u(\cdot + r)) &= y(\cdot - \alpha(u) + r) - \frac{\langle u(\cdot + r), y(\cdot - \alpha(u) + r)\rangle}{\langle u(\cdot + r), \partial_x^2 \phi(\cdot - \alpha(u) + r)\rangle} \partial_x^2 \varphi(\cdot - \alpha(u) + r) \\ &= y(\cdot - \alpha(u) + r) - \frac{\langle u, y(-\alpha(u))\rangle}{\langle u, \partial_x^2 \phi(\cdot - \alpha(u))\rangle} \partial_x^2 \phi(\cdot - \alpha(u) + r) \\ &= (Bu)(\cdot + r), \end{aligned}$$

where in the above calculation we have repeatedly used (ii) of lemma 4.1. ■

COROLLARY 4.1. *The solution $u_\lambda = R(\lambda, v)$ of the initial-value problem*

$$du_\lambda/d\lambda = B(u_\lambda),\ u_0 = v,$$

has the following properties:

(i) *R is a C^1 function for $|\lambda| < \lambda_0(v)$ for any $v \in U_\epsilon$,*
(ii) *R commutes with translations for each λ,*
(iii) *$V(R(\lambda, v))$ is independent of λ, and*
(iv) $\partial R/\partial\lambda(0, \phi) = y$.

Proof. Conclusions (i), (ii) and (iv) are obvious from the properties of B given in the lemma. For (iii), simply note that

$$\mathrm{d}V(u_\lambda)/\mathrm{d}\lambda = \langle u_\lambda, \mathrm{d}u_\lambda/\mathrm{d}\lambda\rangle = \langle u_\lambda, B(u_\lambda)\rangle = 0. \qquad \blacksquare$$

The next three lemmas follow very closely the abstract results of (Grillakis *et al.* (1987)) and consequently their proofs are somewhat abbreviated.

LEMMA 4.2. *There is a C^1 functional $\Lambda:\{v \in U_\epsilon: V(v) = V(\phi)\} \to \mathbb{R}$ such that*

$$E(R(\Lambda(v), v)) > E(\phi)$$

for all $v \in U_\epsilon$ which are not translates of ϕ and are such that $V(v) = V(\varphi)$.

Proof. Let $G(u) = u(\cdot - \alpha(u))$ and solve the equation

$$\langle G(R(\lambda, v)) - \phi, \chi\rangle = 0, \tag{4.3}$$

locally near $\lambda = 0$ and $v = \phi$ by the implicit function theorem. Then $G(u_\lambda) - \phi$ is perpendicular to both χ and $\partial_x \phi$, and so it belongs to the positive subspace of $\mathscr{L}$. By Taylor's theorem,

$$E(u_\lambda) = E(\phi) + \tfrac{1}{2}\langle \mathscr{L}(G(u_\lambda) - \phi), G(u_\lambda) - \phi\rangle + O(\|G(u_\lambda) - \phi\|^2),$$

and so for λ small enough,

$$E(u_\lambda) > E(\phi),$$

unless $G(u_\lambda) = \varphi$. The result follows. (For more details see lemma 4.3 of (Grillakis *et al.* (1987).) $\blacksquare$

LEMMA 4.3. *For $v \in U_\epsilon$ with $V(u) = V(\phi)$ and v not a translate of ϕ, we have*

$$E(\phi) < E(v) + \Lambda(v)\langle E'(v), B(v)\rangle.$$

Proof. This follows from lemma 4.1 and Taylor's theorem. $\blacksquare$

LEMMA 4.4. *The curve ψ_ω constructed in* §3 *satisfies $E(\psi_\omega) < E(\phi)$ for $\omega \neq c$, $V(\psi_\omega) = V(\phi)$ and $\langle E'(\psi_\omega), B(\psi_\omega)\rangle$ changes sign as ω passes through c.*

Proof. Apply lemma 4.3 with $v = \psi_\omega$ to derive that

$$\langle E'(\psi_\omega), B(\psi_\omega)\rangle \Lambda(\psi_\omega) > 0, \tag{4.4}$$

for $\omega \neq c$. So it suffices to show that $\Lambda(\psi_\omega)$ changes sign. Indeed, setting $u_\lambda = R(\Lambda(\psi_\omega), \psi_\omega)$ in (4.3) and differentiating with respect to ω gives

$$\left\langle G'(u_\lambda)\left[\frac{\partial R}{\partial \lambda}\frac{\mathrm{d}\Lambda(\psi_\omega)}{\mathrm{d}\omega} + \frac{\partial R}{\partial v}\frac{\mathrm{d}\psi_\omega}{\mathrm{d}\omega}\right], \chi\right\rangle = 0.$$

Letting $\omega = c$ leads to the formula

$$\langle G'(\phi)\,[y\,\mathrm{d}\Lambda(\psi_\omega)/\mathrm{d}\omega|_{\omega=c} + y], \chi\rangle = 0.$$

But

$$\langle G'(\phi)\,y, \chi\rangle = \langle y, \chi\rangle \neq 0.$$

and therefore

$$\mathrm{d}\Lambda(\psi_\omega)/\mathrm{d}\omega|_{\omega=c} = -1 \neq 0.$$

This concludes the proof of the lemma. $\blacksquare$

With these preliminary results in hand, we can mount a direct attack on the instability portion of theorem 4.1. A particular and somewhat unusual difficulty encountered is that the Lyapunov functional arising naturally in our proof is not obviously finite, and even if it is, its value on a solution path grows with time. An appropriate limitation on its rate of growth is a crucial aspect of the argument presented below.

Proof of instability. Let $\epsilon > 0$ be given and let U_ϵ be the neighbourhood of the orbit of ϕ defined earlier. Without loss of generality, ϵ is supposed to be small enough that lemma 4.1 and its consequences apply within U_ϵ. To demonstrate the instability of ϕ, it suffices to show that there are elements $u_0 \in X$ that are arbitrarily close to ϕ, but for which the solution u of (1.1) with initial data u_0 exits from U_ϵ in finite time. For a fixed u_0, let $[0, t_1)$ denote the maximal interval over which $u(\cdot, t)$ lies continuously in U_ϵ, which is a bounded subset $H^{\frac{1}{2}\mu}(\mathbb{R})$. By theorem 2.1, either $t_1 < t_*$ or $t_1 = \infty$. Our purpose now is to show there are u_0 arbitrarily close to ϕ for which $t_1 < \infty$.

By lemma 4.4 there are elements $u_0 \in X$ arbitrarily close to ϕ that are not translates of ϕ, and are such that $V(u_0) = V(\phi)$, $E(u_0) < E(\phi)$ and $\langle E'(u_0), B(u_0)\rangle > 0$. In view of (3.9), (3.10), and the construction of the curve ψ_ω given earlier, we may assume that $u_0 \in H^{1+\mu}(\mathbb{R})$ and $\int_{-\infty}^{\infty} (1+|x|)^{\frac{1}{2}} |u_0(x)| \, dx < \infty$. In view of the results expounded in §2, the solution u from u_0 enjoys the following properties:

$$\left.\begin{array}{l} u \in C([0, t_1); H^{\mu+1}), \\ E(u(\cdot, t)) \quad \text{and} \quad V(u(\cdot, t)) \quad \text{are constant, for} \quad 0 \leqslant t < t_1, \\ I(u(\cdot, t)) \quad \text{converges and is constant for} \quad 0 \leqslant t < t_1, \\ \sup\limits_{z_1, z_2 \in \mathbb{R}} \left| \int_{z_1}^{z_2} u(x, t) \, dx \right| \leqslant c(1 + t^{\mu/(1+\mu)}), \\ \sup\limits_{0 \leqslant t < t_1} \|u(\cdot, t)\|_{H^{\frac{1}{2}\mu}} \leqslant C, \end{array}\right\} \tag{4.5}$$

where C depends only on ϕ and ϵ, and c depends only on C and $\int_{-\infty}^{\infty} (1+|x|) |u_0(x)| \, dx$.

Let $\beta(t) = \alpha(u(t))$, $Y(x) = \int_{-\infty}^{x} y(z) \, dz$, and define

$$A(t) = \int_{-\infty}^{\infty} Y(x - \beta(t)) \, u(x, t) \, dx. \tag{4.6}$$

The function A serves as a Lyapunov function in our argument. The integral in (4.6) converges. Indeed, if $\mathscr{H}$ is the Heaviside function and $\gamma = \int_{-\infty}^{\infty} y(x) \, dx$, then

$$A(t) = \int_{-\infty}^{\infty} [Y(x - \beta(t)) - \gamma \mathscr{H}(x - \beta(t))] \, u(x, t) \, dx + \gamma \int_{\beta(t)}^{\infty} u(x, t) \, dx.$$

It follows therefore that

$$|A(t)| \leqslant \|Y - \gamma \mathscr{H}\|_{L^2(\mathbb{R})} \|u(t)\|_{L^2(\mathbb{R})} + C'(1 + t^{\mu/(1+\mu)}),$$

and consequently

$$|A(t)| \leqslant C''(1 + t^{\mu/(1+\mu)}) \tag{4.7}$$

for some constant C'', because $V(u(t))$ is constant and $Y-\gamma\mathscr{H} \in L^2(\mathbb{R})$. The latter assertion is a consequence of the following argument. In the light of (3.9), (3.10) and (3.16), it is observed that $(1+|x|)\, y \in L^1(\mathbb{R})$. It follows from Minkowski's inequality that

$$\left[\int_{-\infty}^{0} (Y(x)-\gamma\mathscr{H}(x))^2 \,dx\right]^{\frac{1}{2}} = \left\{\int_{-\infty}^{0} Y^2(x)\,dx\right\}^{\frac{1}{2}} = \left\{\int_{-\infty}^{0}\left(\int_{-\infty}^{x} y(\xi)\,d\xi\right)^2 dx\right\}^{\frac{1}{2}}$$
$$\leqslant \int_{-\infty}^{0}\left(\int_{\xi}^{0} y^2(x)\,dx\right)^{\frac{1}{2}} d\xi = \int_{-\infty}^{0} \sqrt{\xi}\,|y(\xi)|\,d\xi < \infty,$$

Similarly, for $x > 0$,

$$Y(x)-\gamma\mathscr{H}(x) = \int_{x}^{+\infty} y(\xi)\,d\xi,$$

and an analogous inequality holds. Thus $Y-\gamma\mathscr{H} \in L^2(\mathbb{R})$.

Having established that $A(t)$ is well defined and satisfies (4.7), its derivative with respect to time becomes an object of study. The previous formulae obtained in this section make the calculation and estimation of dA/dt relatively straightforward. First of all,

$$\frac{dA}{dt} = -\beta'(t)\int y(x-\beta(t))\,u(x,t)\,dx + \int Y(x-\beta(t))\frac{\partial u}{\partial t}(x,t)\,dx.$$

Note that
$$\beta'(t) = d\alpha(u(t))/dt = \langle \alpha'(u), \partial u/\partial t\rangle,$$
so that
$$dA/dt = \langle -\langle y(\cdot-\beta), u\rangle\, \alpha'(u) + Y(\cdot-\beta), \partial u/\partial t\rangle.$$
Because
$$\partial u/\partial t = \partial_x(Mu-u-f(u)) = \partial_x E'(u),$$
it follows that
$$dA/dt = \langle\langle y(\cdot-\beta, u\rangle\, \partial_x \alpha'(u) - y(\cdot-\beta), E'(u)\rangle,$$
after an integration by parts. The definition of B thus leads to the compact formula

$$dA/dt = \langle B(u), E'(u)\rangle. \tag{4.8}$$

Because
$$0 < E(\phi)-E(u_0) = E(\phi)-E(u(t)),$$
lemma 4.3 implies that
$$0 < A(u(t))\,\langle E'(u(t)), B(u(t))\rangle.$$

Because $u(t) \in U_\epsilon$ and $A(\phi) = 0$, it may be assumed that $A(u(t)) < 1$ by choosing ϵ smaller if necessary. Therefore for all t in $[0, t_1)$,

$$\langle E'(u(t)), B(u(t))\rangle \geqslant E(\phi)-E(u_0) > 0.$$

Hence (4.8) yields the lower bound.

$$dA/dt \geqslant E(\phi)-E(u_0) > 0 \quad \text{for all} \quad t \in [0, t_1). \tag{4.9}$$

Comparing (4.7) and (4.9), it is concluded that $t_1 < \infty$, which means that $u(\cdot, t)$ eventually leaves the tube U_ϵ. This implies instability, and completes one half of the proof of the theorem. ∎

5. Stability

Complementing the theory of instability presented in §4 is the stability theory presented in the present section. The stability of solitary-wave solutions of (1.1) is an immediate consequence of the fact that $d''(c) > 0$ implies that ϕ_c is a local minimum of E subject to the constancy of V. This is a general fact, not special to the equations under consideration in this paper. The proof is, in essence, a special case of the theory given by Grillakis *et al.* (1987).

Lemma 5.1. *Let $d''(c) > 0$. If $y \in X$ is orthogonal to both ϕ and $\partial_x \phi$, then* $\langle \mathscr{L}y, y\rangle > 0$.

Proof. Formula (3.13) insures that $d'(c) = V(\phi_c)$, and hence that $0 < d''(c) = \langle \phi, \mathrm{d}\phi/\mathrm{dc}\rangle = -\langle \mathscr{L}\,\mathrm{d}\phi/\mathrm{dc}, \mathrm{d}\phi/\mathrm{dc}\rangle$. Write $\mathrm{d}\phi/\mathrm{dc} = a_0\chi + b_0\partial_x\phi + p_0$, where p_0 is in the positive subspace of $\mathscr{L}$. Recall that $\mathscr{L}\chi = -\lambda^2\chi$ with $\lambda > 0$ and $\mathscr{L}(\partial_x\phi) = 0$. It follows that

$$\langle \mathscr{L}p_0, p_0\rangle < a_0^2\lambda^2.$$

Now suppose that $\langle y, \phi\rangle = \langle y, \partial_x\phi\rangle = 0$ and decompose y into the sum $a\chi + p$ with p in the positive subspace of $\mathscr{L}$. Because

$$0 = -\langle \phi, y\rangle = \langle \mathscr{L}\,\mathrm{d}\phi/\mathrm{dc}, y\rangle = -a_0 a\lambda^2 + \langle \mathscr{L}p_0, p\rangle,$$

it is inferred that

$$\begin{aligned}\langle \mathscr{L}y, y\rangle &= -a^2\lambda^2 + \langle \mathscr{L}p, p\rangle \geqslant -a^2\lambda^2 + \langle \mathscr{L}p, p_0\rangle^2/\langle \mathscr{L}p_0, p_0\rangle \\ &> -a^2\lambda^2 + (a_0 a\lambda^2)^2/a_0^2\lambda^2 = 0,\end{aligned}$$

as required. ■

Lemma 5.2. *Let $d''(c) > 0$. There exist constants $C > 0$ and $\epsilon > 0$ such that*

$$E(u) - E(\phi) \geqslant c\|u(\cdot + \alpha(u)) - \varphi\|^2_{H^{\frac{1}{2}\mu}},$$

for all $u \in U_\epsilon$ which satisfy $V(u) = V(\phi)$.

Proof. Write u in the form $u(\cdot + \alpha(u)) = (1+a)\phi + y$ where $\langle y, \phi\rangle = 0$ and a is a scalar. Then, by the translation invariance of V and Taylor's theorem,

$$\begin{aligned}V(\phi) = V(u) &= V(u(\cdot + \alpha(u)) \\ &= V(\phi) + \langle \phi, u(\cdot + \alpha(u)) - \phi\rangle + O(\|u(\cdot + \alpha(u)) - \phi\|^2).\end{aligned}$$

Here, and throughout this section the norm is that of $X = H^{\frac{1}{2}\mu}(\mathbb{R})$. The middle term is precisely $a\|\phi\|^2$, so that

$$a = O(\|u(\cdot + \alpha(u)) - \phi\|^2).$$

Writing $L = E + cV$, another Taylor expansion gives

$$L(u) = L(u(\cdot + \alpha(u))) = L(\phi) + \tfrac{1}{2}\langle \mathscr{L}v, v\rangle + o(\|v\|^2),$$

where $v = u(\cdot + \alpha(u)) - \phi = a\phi + y$. This can be written as

$$\begin{aligned} E(u) - E(\phi) &= \tfrac{1}{2}\langle \mathscr{L}v, v\rangle + o(\|v\|^2) \\ &= \tfrac{1}{2}\langle \mathscr{L}y, y\rangle + O(a^2) + O(a\|v\|) + o(\|v\|^2) \\ &= \tfrac{1}{2}\langle \mathscr{L}y, y\rangle + o(\|v\|^2). \end{aligned}$$

But y is orthogonal to both ϕ and $\partial_x \phi$. Therefore by lemma 5.1

$$E(u) - E(\phi) \geqslant 2C\|y\|^2 + o(\|v\|^2)$$

for some constant C. Because

$$\|y\| = \|v - a\varphi\| \geqslant \|v\| - O(\|v\|^2),$$

for $\|v\|$ small, it follows that

$$E(u) - E(\phi) \geqslant C\|v\|^2. \qquad \blacksquare$$

The proof of theorem 4.1 is now completed by showing that $d''(c) > 0$ implies the associated solitary wave to be stable in the sense of definition 4.1.

Proof of stability. Assume $d''(c) > 0$. Let $u_n^0 \in X$ be any sequence such that

$$\inf_s \|u_n^0 - \phi(\cdot + s)\| \to 0 \quad \text{as} \quad n \to \infty.$$

If u_n is the unique solution of (1) with initial data u_n^0, let $\{t_n\}$ be an arbitrary sequence of times such that, for each $n, u_n(\cdot, t_n) \in \partial U_{\frac{1}{2}\epsilon}$. Because E and V are continuous on X and translation invariant,

$$E(u_n(\cdot, t_n)) = E(u_n^0) \to E(\phi) \quad \text{and} \quad V(u_n(\cdot, t_n)) = V(u_n^0) \to V(\phi).$$

Next choose $w_n \in U_\epsilon$ so that $V(w_n) = V(\phi)$ and $\|w_n - u_n(\cdot, t_n)\| \to 0$. By lemma 5.2,

$$0 \leftarrow E(w_n) - E(\phi) \geqslant C\|w_n(\cdot + \alpha(w_n)) - \phi\|^2 = C\|w_n - \phi(\cdot - \alpha(w_n))\|^2,$$

and therefore

$$\|u_n(\cdot, t_n) - \varphi(\cdot - \alpha(w_n))\| \to 0.$$

This means that $u_n(\cdot, t_n)$ tends to the orbit of ϕ. This contradiction proves that the orbit of ϕ is stable.

Putting together the result of the previous section, the fact that the set $\{c > 1 : \phi_c \text{ is stable}\}$ is open, and the present developments, theorem 4.1 is established. $\blacksquare$

Part of this work was done while we were visiting the Institute for Mathematics and its Applications, University of Minnesota. J.L.B. was partly supported by the NSF under grant no. DMS 82-03338. P.E.S. was partly supported by the NSF under grants DMS 84-01725 and DMS 86-01258 and the ONR under contract no. N00014-83-K-0542. W.A.S. was partly supported by the NSF under grant no. DMS 84-20957.

References

Abdelouhab, L., Bona, J. L., Felland, M. & Saut, J.-C. 1987 Non-local models for nonlinear, dispersive waves. *Physica* D (In the press.)

Albert, J. P., Bona, J. L. & Felland, M. 1987 On blow-up for nonlinear, dispersive wave equations. *Mat. Apl. Computação.* (In the press.)

Albert, J. P., Bona, J. L. & Henry, D. 1987 Sufficient conditions for stability of solitary-wave solutions of model equations for long waves. *Physica* D (In the press.)

Benjamin, T. B. 1967 Internal waves of permanent form in fluids of great depth. *J. Fluid Mech.* **29**, 559–592.

Benjamin, T. B. 1972 The stability of solitary waves. *Proc. R. Soc. Lond.* A **328**, 153–183.

Bennett, D. P., Brown, S. E., Stansfield, J. D., Stroughair, J. D. & Bona, J. L. 1983 The stability of internal solitary waves. *Proc. Camb. phil. Soc.* **94**, 351–379.

Bona, J. L. 1975 On the stability of solitary waves. *Proc. R. Soc. Lond.* A **344**, 363–374.

Bona, J. L., Dougalis, V. A. & Karakashian, O. A. 1986 Fully discrete Galerkin methods for the Korteweg–de Vries equation. *Computat Math. Appl.* **12A**, 859–884.

Eckhaus, W. & Schuur, P. 1983 The emergence of solitons of the Korteweg–de Vries equation from arbitrary initial conditions. *Math. Methods Appl. Sci.* **5**, 97–116.

Fornberg, B. & Whitham, G. 1978 A numerical and theoretical study of certain nonlinear wave phenomena. *Phil. Trans. R. Soc. Lond.* A **289**, 373–404.

Gardner, C. S., Greene, J. M., Kruskal, M. D. & Miura, R. M. 1967 A method for solving the Korteweg–de Vries equation. *Phys. Rev. Lett.* **19**, 1095–1097.

Grillakis, M., Shatah, J. & Strauss, W. A. 1987 Stability theory of solitary waves in the presence of symmetry. *J. Funct. Anal.* (In the press.)

Korteweg, D. J. & de Vries, G. 1895 On the change of form of long waves advancing in a rectangular canal, and on a new type of long stationary waves. *Phil. Mag.* **39** (5), 422–443.

McLeod, J. B. & Olver, P. J. 1983 The connection between partial differential equations solvable by inverse scattering and ordinary differential equation of Painleve type. *SIAM Jl math. Anal.* **14**, 488–506.

Scharf, G. & Wreszinski, W. F. 1981 Stability for the Korteweg–de Vries equation by inverse scattering theory. *Ann. Physics* **134**, 56–75.

Shatah, J. 1983 Stable standing waves of nonlinear Klein–Gordon equations. *Communs math. Phys.* **91**, 313–327.

Shatah, J. & Strauss, W. A. 1985 Instability of nonlinear bound states. *Communs math. Phys.* **100**, 173–190.

Weinstein, M. 1986 Lyapunov stability of ground states of nonlinear dispersive evolution equations. *Communs pure appl. Math.* **39**, 51–68.