Proc. R. Soc. Lond. A. **344**, 363–374 (1975)
Printed in Great Britain

On the stability theory of solitary waves

By J. Bona

Department of Mathematics, The University of Chicago and
Fluid Mechanics Research Institute, University of Essex

(*Communicated by T. B. Benjamin, F.R.S. – Received* 18 *November* 1974)

Improvements are made on the theory for the stability of solitary waves developed by T. B. Benjamin. The results apply equally to the Korteweg–de Vries equation and to an alternative model equation for the propagation of long waves in nonlinear dispersive media.

1. Introduction

A theory is developed relating to the stability of solitary-wave solutions of the Korteweg–de Vries equation and of an alternative model equation for unidirectional propagation of long waves. The analysis contained herein is inspired by the work of Benjamin (1972) and attempts to improve upon his results in several aspects.

The accomplishment of the above mentioned paper is considerable, giving a precise formulation and proof of the stability of the shape of solitary-wave solutions of all amplitudes of the Korteweg–de Vries equation. Benjamin's proof deals with the full nonlinear problem, calling upon interesting ideas from functional analysis combined with refined estimates of certain integrals made by use of spectral theory. The proof leads therefore to a much more satisfactory theory than the linearized perturbation analysis of the same problem given earlier by Jeffrey & Kakutani (1970).

The methods used here simply exploit the machinery erected by Benjamin, no essentially new ideas being required. Nevertheless the results presented here are considerably sharper and more satisfactory than those given by Benjamin. The improvement on Benjamin's results takes essentially three forms. First, it is assumed by Benjamin that *solutions* to the model equations under consideration are C^{∞} functions all of whose derivatives decrease sufficiently rapidly to zero at infinity. This hypothesis is replaced by the much more modest assumption that the *initial data* is an L_2 function whose first two derivatives are also in L_2. Secondly, an error in inference is corrected by the derivation of an additional *a priori* bound. Lastly, an *ad hoc* conjecture made by Benjamin near the culmination of his proof is replaced by inequalities which imply the desired result. This latter is important, for the conjecture is not established by Benjamin, and indeed it appears on careful consideration that the conjecture itself is unlikely to be valid. The proof as conceived

After a few preliminary remarks in § 2, Benjamin's theory is briefly outlined in § 3. The refinements of the Benjamin theory are then given precise formulation and proof in the final section.

2. Notation and preliminary remarks

Consideration is given by Benjamin principally to the stability of solitary-wave solutions of the Korteweg–de Vries (1895) equation

$$Nu = u_t + u_x + uu_x + u_{xxx} = 0, \tag{2.1}$$

though solitary-wave solutions of the alternative model equation

$$\tilde{N}u = u_t + u_x + uu_x - u_{xxt} = 0, \tag{2.2}$$

proposed by Benjamin, Bona & Mahony (1972) are also shown to be stable by an almost identical analysis. For the Korteweg–de Vries (K.–dV.) equation the solitary-wave solution is written explicitly as $u_s(x,t) = \phi(x-\bar{C}t)$ where $\bar{C} = 1+C$, $C > 0$, and

$$\phi(z) = 3C\,\mathrm{sech}^2\,(\tfrac{1}{2}C^{\frac{1}{2}}z), \tag{2.3}$$

while for the equation (2.2) the solitary wave takes the form $\tilde{u}_s(x,t) = \tilde{\phi}(x-\bar{C}t)$ with $\bar{C}$ as before and

$$\tilde{\phi}(z) = 3C\,\mathrm{sech}^2\,[C^{\frac{1}{2}}z/2(1+C)^{\frac{1}{2}}]. \tag{2.4}$$

Of course, an arbitrary translation of ϕ or $\tilde{\phi}$ may also be used to define a solution of the respective equation.

The stability problem in question refers to the pure initial-value problem for (2.1) or (2.2) posed on the domain $\{(x,t)\colon x\in\mathbb{R},\, t\geqslant 0\}$:

$$\left.\begin{aligned} Nu = 0, \quad u(x,0) = \psi(x), \\ \text{and} \qquad \tilde{N}\tilde{u} = 0, \quad \tilde{u}(x,0) = \tilde{\psi}(x). \end{aligned}\right\} \tag{2.5}$$

Suppose that ψ is close enough to ϕ (respectively $\tilde{\psi}$ is close enough to $\tilde{\phi}$) in a certain sense. The conclusion in view is that the solution u (respectively $\tilde{u}$) of (2.5) is then close to u_s (respectively $\tilde{u}_s$), in an appropriate sense, for all $t \geqslant 0$. To give precision to the above statement, there must be defined measures of distance between elements of various function classes.

Let L_2 be the measurable functions from $\mathbb{R}$ to $\mathbb{R}$ which are square-integrable. L_2 is given its usual norm, denoted by $\|\cdot\|$. Let k be a non-negative integer and let H^k denote the Sobolev space of L_2 functions whose (generalized) derivatives up to order k are in L_2. This space is given its standard Hilbert space structure with norm

$$\|f\|_k^2 = \sum_{j=0}^{k} \|f^{(j)}\|^2.$$

For k as above, and T either a finite positive number or $+\infty$, let $\mathscr{H}_T^k = C(0,T;H^k)$ be the functions $u\colon \mathbb{R}\times[0,T]\to\mathbb{R}$ such that $u(\cdot,t)\in H^k$ for each t and such that the

As pointed out by Benjamin *et al.* (1972) it is not true that, because an initial wave profile lies close to the profile of a particular solitary wave, the corresponding solution of K.–dV. or (2.2) will continue to lie close to the solitary wave in question as it evolves in time. This owes to the fact that the solitary wave and the perturbation may have different speeds of propagation. The proposed global stability result therefore takes no notice of the spatial placement of the waves in question. Mathematically, this is reflected in the introduction of a 'sliding' metric. For $f, g \in H^1$ define

$$d(f,g) = \inf_{y \in \mathbb{R}} \|f(\cdot + y) - g(\cdot)\|_1. \tag{2.6}$$

Then d is a pseudo-metric on H^1 and defines a proper metric on the quotient space H^1/G, where G is the translation group in $\mathbb{R}$. $d(f,g)$ being small means that a translation of f lies close to g in H^1 norm. But if $h \in H^1$, h is a bounded and continuous function with

$$\sup_{x \in \mathbb{R}} |h(x)| \leqslant \|h\|_1, \tag{2.7}$$

so, if $d(f,g)$ is small, it is further implied that a translation of f lies pointwise uniformly near to g: that is, f and g have nearly the same shape. The stability theorem formulated by Benjamin, in the notation of (2.5), says essentially that if $\psi - \phi$ (respectively $\tilde{\psi} - \tilde{\phi}$) is small in H^1 norm then $d(\phi, u)$ (respectively $d(\tilde{\phi}, \tilde{u})$) is small for all $t \geqslant 0$, so that for all later time, u (respectively $\tilde{u}$) has practically the same shape as the original solitary wave to which it was compared at $t = 0$.

The qualitative results for the initial-value problems (2.5) will be needed in the development of the present theory. These are stated below for convenient reference. Proofs may be found in Bona & Smith (1975).

Proposition 1. Let $\psi \in H^s$ where $s \geqslant 1$. Then there exists a solution u in

$$L_\infty(0, \infty; H^s)$$

to the initial-value problem for K.–dV. If $s \geqslant 2$, u is unique, $u \in \mathscr{H}^s_\infty$ and $\partial_t^l u \in \mathscr{H}^{s-3l}_\infty$ for l such that $s - 3l \geqslant 0$.

Let ψ_1 and ψ_2 be elements of H^s where $s \geqslant 2$ and let u_1 and u_2 be respectively the associated solutions of the initial-value problem for K.–dV. Let $T > 0$ be given. Then there exist constants M_k, $k = 0, 1, \ldots s$, such that

$$\|u_1 - u_2\|_k \leqslant M_k \|\psi_1 - \psi_2\|_k \tag{2.8}$$

for all t in $[0, T]$. The constants M_k depend on T, $\|\psi_1\|_k$ and $\|\psi_2\|_k$ if $k \geqslant 2$, and on T, $\|\psi_1\|_2$ and $\|\psi_2\|_2$ if $k = 0$ or 1.

If $s \geqslant 2$ and ψ and u are as above, then

$$V(u) = \int_{-\infty}^{\infty} u^2(x,t)\,\mathrm{d}x \quad \text{and} \quad M(u) = \int_{-\infty}^{\infty} (u_x^2 - \tfrac{1}{3}u^3)\,\mathrm{d}x \tag{2.9}$$

PROPOSITION 2. Let $\psi \in H^s$ where $s \geqslant 1$. Then there exists a unique solution u to the initial-value problem for (2.2) which lies in $\mathscr{H}^1_\infty$ and for each $T > 0$, in $\mathscr{H}^s_T$. Further for all $l > 0$ $\partial^l_t u$ lies in $\mathscr{H}^1_\infty$ and, for each $T > 0$, in $\mathscr{H}^{s+1}_T$.

Let ψ_1 and ψ_2 be elements of H^s where $s \geqslant 1$ and let u_1 and u_2 respectively be the associated solutions of the initial-value problem for (2.2). Let $T > 0$ be given. Then there exists constants N_k, $k = 0, 1, \ldots, s$ such that

$$\|u_1 - u_2\|_k \leqslant N_k \|\psi_1 - \psi_2\|_k \tag{2.10}$$

for all t in $[0, T]$. The constants N_k depend on T, $\|\psi_1\|_k$ and $\|\psi_2\|_k$ for $k \geqslant 1$ and N_0 depends on T, $\|\psi_1\|_1$ and $\|\psi_2\|_1$.

If $s \geqslant 1$ and ψ and u are as above, then both

$$E(u) = \int_{-\infty}^{\infty} (u^2 + u_x^2)\,\mathrm{d}x \quad \text{and} \quad M(u) = \int_{-\infty}^{\infty} (u_x^2 - \tfrac{1}{3}u^3)\,\mathrm{d}x \tag{2.11}$$

are independent of $t \geqslant 0$.

Remark. There is a sequence of integrals of polynomials in u and its spatial derivatives which, for sufficiently smooth solutions of K.–dV., do not vary with time (cf. Miura, Gardner & Kruskal (1968); Kruskal, Miura, Gardner & Zabusky (1970) and Bona & Smith (1973)).

The following theory applies equally well to K.–dV. or (2.2) with only minor differences in detail. The results will be proven in detail for K.–dV. with the modifications needed for (2.2) then briefly indicated.

3. Outline of Benjamin's theory

Benjamin begins by considering the solution u of the initial-value problem (2.5) for K.–dV. with initial data denoted here by ψ. ψ will be assumed for the moment to lie in $H^\infty = \bigcap_{k=1}^{\infty} H^k$. It is further supposed that $V(\psi) = V(\phi)$, an extra side-condition which is easily dispensed with later. Define

$$h(x,t) = u(x,t) - \phi(x+a), \tag{3.1}$$

where $a = a(t)$ is to be chosen subsequently. Since $V(u) = V(\psi) = V(\phi)$ for all $t \geqslant 0$ from proposition 1,

$$\int_{-\infty}^{\infty} (2\phi h + h^2)\,\mathrm{d}x = 0, \tag{3.2}$$

so it appears that

$$\Delta M = \Delta M(\phi, h) = M(u) - M(\phi) = \int_{-\infty}^{\infty} [h_x^2 + (C-\phi)h^2 - \tfrac{1}{3}h^3]\,\mathrm{d}x. \tag{3.3}$$

The usefulness of ΔM derives from the fact that it does not vary with time since $M(u)$ is independent of t from proposition 1. It follows that

$$\Delta M \leqslant \int_{-\infty}^{\infty} (h_x^2 + Ch^2)\,\mathrm{d}x + \tfrac{1}{3} \sup_{x \in \mathbb{R}} |h| \int_{-\infty}^{\infty} h^2\,\mathrm{d}x \leqslant m\|h\|_1^2 + \tfrac{1}{3}\|h\|_1^3, \tag{3.4}$$

The difficult part of the proof is concerned with estimating an effective lower bound for ΔM. Benjamin proceeds by breaking h into even and odd parts. For a fixed $t \geqslant 0$, let

$$h(x,t) = f(x) + g(x), \tag{3.5}$$

where $f(x) = f(-x)$ and $g(x) = -g(-x)$. Writing h in this manner allows ΔM to be written as

$$\Delta M = 2\int_0^\infty [f_x^2 + (C-\phi)f^2]\,\mathrm{d}x + 2\int_0^\infty [g_x^2 + (C-\phi)g^2]\,\mathrm{d}x - \tfrac{1}{3}\int_{-\infty}^\infty h^3\,\mathrm{d}x. \tag{3.6}$$

The contribution to the quadratic portion of ΔM from the odd and even portions of h may therefore be estimated separately.

Benjamin shows, by adroit use of the spectral theory of the eigenvalue problem

$$\rho'' + [20\,\mathrm{sech}^2(\chi) + \lambda]\rho = 0,$$

$$\rho'(0) = 0, \quad \rho \text{ bounded on } [0,\infty),$$

that
$$\int_{-\infty}^\infty [f_x^2 + (C-\phi)f^2]\,\mathrm{d}x \geqslant c_1 \int_{-\infty}^\infty (f_x^2 + Cf^2)\,\mathrm{d}x - c_2\|h\|_1^3, \tag{3.7}$$

where c_1 and c_2 are positive constants which depend only on C, which is fixed throughout the discussion by the choice of the initial solitary wave. The inequality (3.7) is derived by using the side condition (3.1) but without having to specify the translation constant a appearing in the definition of h.

Next the contribution to ΔM from the odd part of h is estimated. For this purpose let $a = a(t)$ be defined by

$$\int_{-\infty}^\infty [u(x,t) - \phi(x+a)]^2\,\mathrm{d}x = \inf_{y\in\mathbb{R}} \int_{-\infty}^\infty [u(x,t) - \phi(x+y)]^2\,\mathrm{d}x. \tag{3.8}$$

Benjamin claims at this point that the infimum is attained at a finite value of y simply because u and ϕ decay to zero at ∞. This inference is incorrect as it stands. If $u < 0$ everywhere, for example, the infimum plainly occurs only at $\pm\infty$. It may be argued that as long as u resembles even roughly the solitary wave ϕ the infimum will surely be taken at finite values of y, but this leads to circularity. What is required is some definite means of assuring the infimum is taken on at finite values of y. A method of guaranteeing this will be given in the next section.

On any time interval in which the infimum in (3.8) obtains at finite values, Benjamin's arguments will apply. In particular the definition (3.1) of h will make sense and

$$\|h\|_1 = \|u(\cdot,t) - \phi(\cdot + a)\|_1 \geqslant d(u,\phi) \tag{3.9}$$

by definition of d. On differentiating (3.8) with respect to a, the condition appears. This condition is used, together with another application of spectral theory, this time applied to the eigenvalue problem

$$\theta'' + (\phi(x) + \lambda C)\theta = 0,$$

$$\theta(0) = 0, \quad \theta \text{ bounded on } [0, \infty),$$

to deduce that

$$\int_{-\infty}^{\infty} [g_x^2 + (C-\phi)g^2]\,\mathrm{d}x \geqslant \frac{1}{4}\int_{-\infty}^{\infty} (g_x^2 + Cg^2)\,\mathrm{d}x \geqslant c_3\|g\|_1^2, \tag{3.10}$$

where $c_3 = \frac{1}{4}\min(1, C)$. It follows from (3.7), (3.10) and the estimate

$$\frac{1}{3}\int_{-\infty}^{\infty} h^3\,\mathrm{d}x \leqslant \tfrac{1}{3}\|h\|_1 \int_{-\infty}^{\infty} h^2\,\mathrm{d}x$$

that

$$\Delta M \geqslant c_4\|h\|_1^2 - c_5\|h\|_1\|h\|^2 \geqslant c_4\|h\|_1^2 - c_5\|h\|_1^3, \tag{3.11}$$

where c_4 and c_5 are positive constants depending only on C.

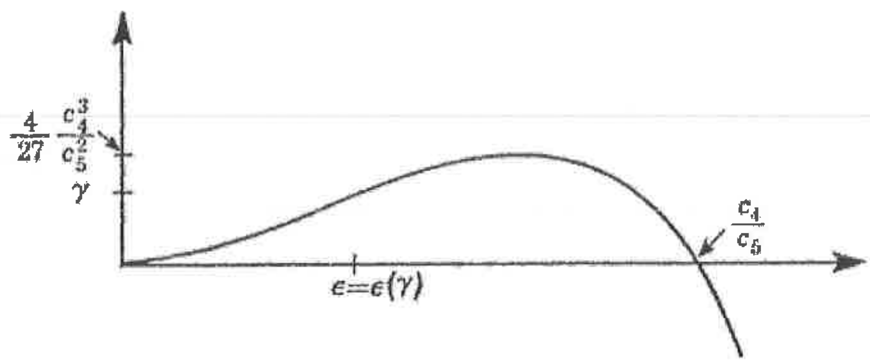

FIGURE 1. Graph of the polynomial in (3.12).

Benjamin now supposes that $\|h\|_1$ is varied continuously in time and that $\|\psi - \phi\|_1 = \delta$ is small. It follows from (3.4) that $\Delta M \leqslant m\delta^2 + \frac{1}{3}\delta^3$ at $t = 0$. Since ΔM is independent of time, the same inequality holds for all $t \geqslant 0$. Combining this with (3.11) yields

$$\gamma(\delta) = m\delta^2 + \tfrac{1}{3}\delta^3 \geqslant c_4\|h\|_1^2 - c_5\|h\|_1^3. \tag{3.12}$$

The cubic polynominal $p(x) = c_4x^2 - c_5x^3$ whose value at $\|h\|_1$ is on the right of (3.12) has a graph as shown in figure 1. Since $\|h\|_1 \geqslant 0$ always, if $\gamma < 4c_4^3/27c_5^2$, and $\|h\|_1$ is a continuous function of t, it follows from (3.12) that

$$\epsilon = \epsilon(\gamma) \geqslant \|h\|_1,$$

where $\epsilon(\gamma)$ is the smallest positive root of

$$p(\epsilon) = \gamma.$$

ϵ may be made as small as desired by taking γ small, and from (3.12), γ is small for δ small enough. (3.9) thus implies that for all $t \geqslant 0$,

$$\epsilon \geqslant d(u, \phi)$$

The problem with this argument is that $\|h_1\|$ must be assumed to be a continuous function of t. This is not established by Benjamin, and moreover seems unlikely to be true because, according to (3.8), the constant $a = a(t)$, and therefore h itself, is defined only in terms of the L_2 structure of u and ϕ. It will follow, as in the next section, that $\|h\|$ is a continuous function of time, but there is nothing in the determination of h which would keep $\|h\|_1$ from shifting discontinuously in time by means of a sudden change in $\|h_x\|$. This is a point which will be resolved in the next section.

Finally the assumption that $V(\psi) = V(\phi)$ is removed by a simple application of the triangle inequality for the pseudo-metric d. Suppose ψ is arbitrary with $\|\psi - \phi\|_1$ small. Let ϕ_1 be the solitary-wave solution of K.–dV. such that $V(\phi_1) = V(\psi)$. Then $d(u, \phi) \leqslant d(u, \phi_1) + d(\phi_1, \phi)$. Now $d(\phi, \phi_1)$ is a constant which is small if $|V(\phi) - V(\phi_1)|$ is small. The latter is true if $\|\psi - \phi\|$ is small. Moreover, the above theory applies to u viewed as a perturbation of ϕ_1 and shows that $d(u, \phi_1)$ is small provided $\|\psi - \phi_1\|_1$ is small, which is the case since both $\|\psi - \phi\|_1$ and $\|\phi - \phi_1\|_1$ are small.

This completes the argument given by Benjamin. The proof is very elegant and the final result provides a satisfactory solution to the question of the stability of solitary-wave solutions of the two model equations for long waves considered. The next section is devoted to the previously mentioned improvements and corrections of Benjamin's basic argument.

4. Stability of the solitary wave

This section contains the major contribution of the present paper. As in the previous sections, ϕ denotes a solitary-wave solution of K.–dV. (while $\tilde{\phi}$ denotes a solitary-wave solution of (2.2)) with wave speed $\bar{C} = 1 + C$ where $C > 0$.

Theorem 1. Let $\epsilon > 0$ be given. Then there exists $\bar{\delta} > 0$ such that if $\psi \in H^2$, with u the solution of K.–dV. corresponding to the initial data ψ, and $\|\psi - \phi\|_1 \leqslant \bar{\delta}$, then $d(u, \phi) \leqslant \epsilon$ for all $t \geqslant 0$.

Theorem 2. Let $\epsilon > 0$ be given. Then there exists $\bar{\delta} > 0$ such that if $\tilde{\psi} \in H^1$, with $\tilde{u}$ the solution of (2.2) corresponding to the initial data $\tilde{\psi}$, and $\|\tilde{\psi} - \tilde{\phi}\|_1 \leqslant \bar{\delta}$, then $d(\tilde{u}, \tilde{\phi}) \leqslant \epsilon$ for all $t \geqslant 0$.

Remark. The theorem for the model equation (2.2) is slightly better than the result for K.–dV. in that the initial data is only required to lie in H^1. The H^1 result is plainly optimal in the present context. The lack of a continuous dependence result in H^1 is the reason for the failure of the stability theorem for K.–dV. in H^1 (cf. proposition 1 and compare to proposition 2).

Proof. Suppose initially that $\psi \in H^\infty$ so that u and all its partial derivatives are elements of $\mathscr{H}_T^1$ for any $T > 0$. Following Benjamin, suppose also that $V(\psi) = V(\phi)$. The theorem is first established for this very smooth situation. Without loss of

Let $u_s(x,t) = \phi(x-\bar{C}t)$ and $w(x,t) = u(x,t) - u_s(x,t)$. By the continuous dependence result in proposition 1, for given $T_0 > 0$ there exists $\delta_0 = \delta_0(T_0, \phi, \|\psi\|_2) > 0$ such that if $\|\psi - \phi\|_1 \leqslant \delta_0$, then

$$\|w\| < 2V(\phi). \tag{4.1}$$

for t in $[0, T_0]$. (4.1) has an important implication given in the following lemma.

LEMMA 1. Let $v(x,t) = \phi(x - \bar{C}t + b)$ and u a solution of K.–dV. corresponding to initial data ψ. Suppose for some $t_0 \geqslant 0$, $\|u(\cdot, t_0) - v(\cdot, t_0)\| < 2V(\phi)$. Then

$$\inf_{y \in \mathbb{R}} \int_{-\infty}^{\infty} \{u(x,t_0) - \phi(x+y)\}^2 \, dx \tag{4.2}$$

is attained at finite values of y.

Proof. Let

$$\rho(y) = \int_{-\infty}^{\infty} \{u(x,t_0) - \phi(x+y)\}^2 \, dx. \tag{4.3}$$

ρ is a continuous function of y and $\lim_{y \to \pm\infty} \rho(y)$ exists. In fact

$$\lim_{y \to \pm\infty} \rho(y) = \int_{-\infty}^{\infty} u^2(x,t_0) \, dx + \int_{-\infty}^{\infty} \phi^2 \, dx = 2V(\phi). \tag{4.4}$$

But by assumption, at $y_0 = \bar{C}t_0 - b$,

$$\rho(y_0) < 2V(\phi). \tag{4.5}$$

The continuity of ρ coupled with (4.4) and (4.5) imply the desired result.

Because of (4.1) and lemma 1 applied to u_s and u, the infimum

$$\inf_{y \in \mathbb{R}} \int_{-\infty}^{\infty} [u(x,t) - \phi(x+y)]^2 \, dx \tag{4.6}$$

is taken on at finite values of y throughout the interval $[0, T_0]$. Hence the definition

$$h(x,t) = u(x,t) - \phi(x+a), \tag{4.7}$$

where a is a number at which the infimum in (4.6) is taken on, is meaningful. The infimum (4.6) being taken on at finite values means that Benjamin's estimates (3.11) hold on $[0, T_0]$. Thus for t in $[0, T_0]$,

$$\Delta M \geqslant c_4 \|h\|_1^2 - c_5 \|h\|_1 \|h\|^2, \tag{4.8}$$

where again, c_4 and c_5 are positive constants dependent only on C, which is fixed in this discussion. Define

$$A = A(t) = \|h\| \quad \text{and} \quad B = B(t) = \|h_x\|. \tag{4.9}$$

Then (4.8) implies

$$\Delta M \geqslant c_4(A^2 + B^2) - c_5(A+B)A^2. \tag{4.10}$$

LEMMA 2. $$A(t) = \inf_{y \in \mathbb{R}} \left\{ \int_{-\infty}^{\infty} [u(x,t) - \phi(x+y)]^2 \, dx \right\}^{\frac{1}{2}}$$

Proof. If $t_1, t_2 \geqslant 0$, then

$$\begin{aligned}|A(t_1)-A(t_2)| &= \left| \inf_{y\in\mathbb{R}} \left\{\int_{-\infty}^{\infty} [u(x,t_1)-\phi(x+y)]^2\,\mathrm{d}x\right\}^{\frac{1}{2}} - \inf_{z\in\mathbb{R}} \left\{\int_{-\infty}^{\infty} [u(x,t_2)-\phi(x+z)]^2\,\mathrm{d}x\right\}^{\frac{1}{2}}\right| \\ &= \inf_{y\in\mathbb{R}}\inf_{z\in\mathbb{R}} \big|\|u(\cdot,t_1)-\phi(\cdot+y)\| - \|u(\cdot,t_2)-\phi(\cdot+z)\|\big| \\ &\leqslant \inf_{y\in\mathbb{R}}\inf_{z\in\mathbb{R}} \{\|u(\cdot,t_1)-u(\cdot,t_2)\| + \|\phi(\cdot+y)-\phi(\cdot+z)\|\} \\ &= \|u(\cdot,t_1)-u(\cdot,t_2)\|.\end{aligned}$$

The solution u is known to be a continuous mapping of $[0,\infty)$ into L_2 from propositions 1 or 2. Hence the continuity of A is established.

Remark. Let $\mathscr{A} = \{\phi(\cdot+b): b\in\mathbb{R}\}$. $\mathscr{A}$ is a bounded subset of L_2 and $A(t)$ is simply the distance of $u(\cdot,t)$ to $\mathscr{A}$ in L_2. Viewed in this light, the proof of lemma 2 reproduces a well-known fact from metric space theory.

Suppose now that $\|\psi-\phi\|_1 = \delta$, where the restrictions on δ will now be determined. Because of (3.4), for any t such that the definition of h makes sense

$$\Delta M \leqslant m\|h\|_1^2 + \tfrac{1}{3}\|h\|_1^3.$$

At $t=0$ it may therefore be inferred that

$$\Delta M \leqslant m\delta^2 + \tfrac{1}{3}\delta^3 = \gamma(\delta) = \gamma, \tag{4.11}$$

and this inequality holds for all $t \geqslant 0$ for which h is defined since ΔM is invariant in time. Writing out this inequality,

$$\int_{-\infty}^{\infty} [h_x^2 + (C-\phi)h^2 - \tfrac{1}{3}h^3]\,\mathrm{d}x \leqslant \gamma.$$

Hence

$$\left.\begin{aligned} B^2 = \int_{-\infty}^{\infty} h_x^2\,\mathrm{d}x &\leqslant \gamma + \int_{-\infty}^{\infty} [\tfrac{1}{3}h^3 + (\phi - C)h^2]\,\mathrm{d}x \\ &\leqslant \gamma + \tfrac{1}{3}\sup|h| \int_{-\infty}^{\infty} h^2\,\mathrm{d}x + 2C\int_{-\infty}^{\infty} h^2\,\mathrm{d}x \\ &\leqslant \gamma + \tfrac{1}{3}(A+B)A^2 + 2CA^2, \end{aligned}\right\} \tag{4.12}$$

since $\phi \leqslant 3C$ everywhere in $\mathbb{R}$. Solving the quadratic inequality in B above, there is implied

$$B \leqslant \tfrac{1}{2}\{\tfrac{1}{3}A^2 + [\tfrac{1}{9}A^4 + 4(\gamma + \tfrac{1}{3}A^3 + 2CA^2)]^{\frac{1}{2}}\} = F(A). \tag{4.13}$$

By using (4.13) in (4.10),

$$\left.\begin{aligned} \gamma \geqslant \Delta M &\geqslant c_4(A^2+B^2) - c_5[A+F(A)]A^2 \\ &\geqslant A^2[c_4 - c_5F(A)] - c_5A^3 = G(A). \end{aligned}\right\} \tag{4.14}$$

Now $F(0)$ ... such as long as

the graph of G is depicted in figure 2, as may be confirmed by a careful study of the function G on the positive real axis. From lemma 2, A is a continuous function of time. Hence the Liapunov-type argument given by Benjamin, which was outlined near figure 1, may be used to assure that so long as

$$\gamma < G_m, \tag{4.16}$$

then

$$A \leqslant A_\gamma, \tag{4.17}$$

where A_γ is the smallest positive root of $G(A) = \gamma$, independent of $t \geqslant 0$. (N.B.: G_m depends on γ of course, but G_m increases as γ decreases to zero, so the condition (4.16) can be met for sufficiently small γ.)

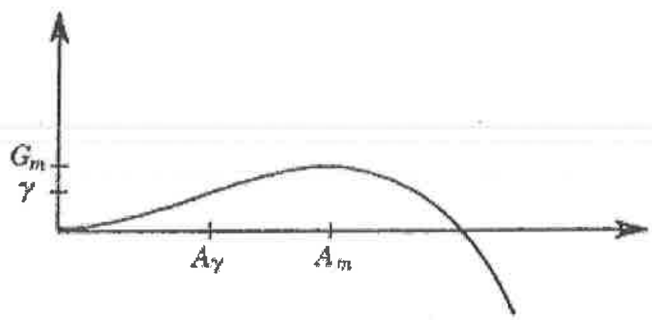

FIGURE 2. Graph of the function G.

With a bound on A in hand, return now to the first line of (4.14) to infer a bound on B.

$$c_4 B^2 \leqslant \gamma + c_5[A + F(A)]A^2 \leqslant \gamma + c_5[A_\gamma + F(A_\gamma)]A_\gamma^2. \tag{4.18}$$

Hence in sum, for γ satisfying (4.15) and (4.16),

$$\|h\|_1^2 = A^2 + B^2 \leqslant A_\gamma^2 + (1/c_4)\{\gamma + c_5[A_\gamma + F(A_\gamma)]A_\gamma^2\}, \tag{4.19}$$

independent of $t \geqslant 0$, so long as the infimum in (4.6) is taken on at finite values of y. By taking δ small, γ as defined in (4.11) may be made small and hence A_γ may be made as near zero as desired. In particular, by taking $\delta \leqslant \delta_1$ say, it follows that

$$A_\gamma^2 + \frac{1}{c_4}\{\gamma + c_5[A_\gamma + F(A_\gamma)]A_\gamma^2\} \leqslant \epsilon^2. \tag{4.20}$$

Let δ_2 be such that if $\delta \leqslant \delta_2$ then (4.15) holds and δ_3 such that if $\delta \leqslant \delta_3$ then (4.16) holds. Let $\bar{\delta} = \min(\delta_0, \delta_1, \delta_2, \delta_3)$. Note that δ_1, δ_2 and δ_3 depend only on ϵ and C while δ_0 depends on ϵ, C and $\|\psi\|_2$ (for K.–dV., and on ϵ, C and $\|\psi\|_1$ for (2.2)). If $\delta \leqslant \bar{\delta}$, then $\|h\|_1^2 \leqslant \epsilon^2$ so long as the infimum in (4.6) is attained at finite values of y, and since $\delta \leqslant \delta_0$, this occurs at least on $[0, T_0]$ from (4.1) and lemma 1. Hence as long as the infimum in (4.6) is attained at finite values, the latter inequality and (3.9) imply

$$d(u, \phi) \leqslant \|h\|_1 \leqslant \epsilon. \tag{4.21}$$

Let $\mathscr{S} = \{t$: the infimum in (4.6) is attained at finite values of $y\}$. $\mathscr{S} \supset [0, T_0)$ from above. Let T_1 be the largest value such that $\mathscr{S} \supset [0, T_1)$ and suppose $T_1 < +\infty$. Then for $t < T_1$

Now A is a continuous function of t for all $t \geqslant 0$. Hence there is a $T > 0$ such that

$$A(t) < 2V(\phi)$$

for t in $[T_1, T_1 + T]$. But then lemma 1 implies the infimum in (4.6) is taken at finite values of y in $[T_1, T_1 + T]$ and this contradicts the choice of T_1. There is left only the conclusion $T_1 = +\infty$ and the stability, conditioned on the auxillary requirement $V(\psi) = V(\phi)$, is established for smooth data.

To remove the condition that $V(\psi) = V(\phi)$ requires only an application of the triangle inequality for d. This has already been outlined at the end of §3.

The removal of the smoothness condition relies on the continuous-dependence result in proposition 1 once more. Specifically for K.–dV., let a solitary wave ϕ, $\epsilon > 0$ and $\psi \in H^2$ be given. Let $\bar{\delta} = \bar{\delta}(\epsilon, \phi, \|\psi\|_2)$ be as provided by the foregoing analysis and suppose $\|\psi - \phi\|_1 < \bar{\delta}$.

Let $\{\psi_n\}_{n=1}^{\infty}$ be a sequence of H^∞ functions such that

$$\left.\begin{aligned} &\psi_n \to \psi \quad \text{in} \quad H^2, \quad \text{as} \quad n \to \infty, \\ &\|\phi - \psi_n\|_1 < \bar{\delta}, \quad \text{for all } n, \\ &\|\psi_n\|_2 \leqslant \|\psi\|_2, \quad \text{for all } n. \end{aligned}\right\} \tag{4.23}$$

The construction of such a sequence presents no problem (cf. Bona & Smith 1975, lemma 5). Let u_n be the solution of K.–dV. associated to the initial data ψ_n, for $n = 1, 2, \ldots$. Then since $\|\phi - \psi_n\|_1 < \bar{\delta}$, for all $t \geqslant 0$

$$d(u_n, \phi) \leqslant \epsilon \tag{4.24}$$

for all n. Note that $\bar{\delta}$ works for ψ_n since $\|\psi_n\|_2 \leqslant \|\psi\|_2$ for all n. The continuous-dependence result of proposition 1 implies that for any finite $T > 0$, $u_n \to u$ in $\mathscr{H}^2_T$. In particular, for each $t \geqslant 0$, $u_n(\cdot, t) \to u(\cdot, t)$ in H^2, and therefore certainly in H^1. It follows that $d(u_n, \phi) \to d(u, \phi)$ for each t. Therefore

$$d(u, \phi) \leqslant \epsilon \tag{4.25}$$

for all $t \geqslant 0$. This completes the proof of the theorem 1.

The proof of theorem 2 follows the same general lines, with the qualitative results in proposition 2 used in the place of the results of proposition 1. Given a solitary wave $\tilde{\phi}$, consider first the special case of initial date $\tilde{\psi} \in H^\infty$ such that $E(\tilde{\psi}) = E(\tilde{\phi})$. Let $\tilde{u}$ be the solution of (2.2) with initial data $\tilde{\psi}$ and let $\tilde{h}(x, t) = \tilde{u}(x,t) - \tilde{\phi}(x + a)$, where a is determined subsequently. Corresponding to (3.3), there appears

$$\Delta M = M(\tilde{u}) - M(\tilde{\phi}) = \int_{-\infty}^{\infty} [(1 + C)\tilde{h}_x^2 + (C - \tilde{\phi})\tilde{h}^2 - \tfrac{1}{3}\tilde{h}^3]\,\mathrm{d}x.$$

Again ΔM does not depend on t. An upper bound on ΔM is available as in (3.4). The estimates from below for ΔM proceed as before by breaking $\tilde{h}$ into even and odd parts and determining bounds on the even and odd part of the quadratic portion of of lemma 1 follows exactly as before while the result (4.1) has an analog that derives from proposition 2. The rest of the reasoning is as above except that proposition 2 replaces the reliance on proposition 1. The removal of the side condition $E(\psi) = E(\phi)$ is effected by the triangle inequality for d as before. Weakening the assumption that $\psi \in H^\infty$ requires another application of the continuous-dependence result in proposition 2. Since for (2.2) continuous dependence holds with regard to H^1 perturbations, the final conclusion is that $d(\tilde{u}, \tilde{\phi}) \leqslant \epsilon$ provided only that $\tilde{\psi} \in H^1$ and $\|\tilde{\psi} - \tilde{\phi}\|_1$ is small enough.

In summary, the stability of the shape of solitary-wave solutions of the Korteweg–de Vries equation is established for H^2 perturbations which are small in H^1 norm. For solitary wave solutions of (2.2), stability in shape is proved for H^1 perturbations which are small in H^1 norm.

The present theory is restricted to the two model equations (2.1) and (2.2). However, questions of the stability of solitary-wave solutions of other one-dimensional model equations for long waves, or indeed for the full equations of surface waves or internal waves in stratified fluid, to name just a few examples, appear to present many of the same sort of problems already encountered for (2.1) or (2.2). Hence the ideas of Benjamin exploited herein may have wider scope than the present context.

The author wishes to thank Professor T. B. Benjamin, F.R.S., for helpful comments on the manuscript. He is indebted to the Science Research Council, U.K. and to the National Science Foundation, U.S.A., for support during portions of time when the present research was carried out.

References

Benjamin, T. B. 1972 The stability of solitary waves. *Proc. R. Soc. Lond.* **A328**, 153.

Benjamin, T. B., Bona, J. L. & Mahony, J. J. 1972 Model equations for long waves in nonlinear dispersive systems. *Phil. Trans. R. Soc. Lond.* **A272**, 47.

Bona, J. L. & Smith, R. 1975 The initial-value problem for the Korteweg–de Vries equation. *Phil. Trans. R. Soc. Lond.* A **278**, 555–601.

Jeffrey, A. & Kakutani, T. 1970 Stability of the Burger's shock wave and the Korteweg–de Vries soliton. *Indiana Univ. Math. J.* **20**, 463.

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.* (5), **39**, 422.

Kruskal, M. D., Miura, R. M., Gardner, C. S. & Zabusky, N. J. 1970 Uniqueness and non-existence of polynomial conservation laws. *J. Math. Phys.* **11**, 952.

Miura, R. M., Gardner, C. S. & Kruskal, M. D. 1968 Existence of conservation laws and constants of motion. *J. Math. Phys.* **9**, 1204.