MODEL EQUATIONS FOR LONG WAVES IN NONLINEAR DISPERSIVE SYSTEMS

By T. B. BENJAMIN, F.R.S., J. L. BONA and J. J. MAHONY†
Fluid Mechanics Research Institute, University of Essex, Colchester, Essex

(*Received* 18 *August* 1971)

CONTENTS

	PAGE
1. Introduction	48
2. Discussion of approximate equations for the evolution of long waves	50
Shortcomings of the KdV equation	54
3. Existence theory for the regularized equation	57
Comments concerning function spaces	62
The effect of a forcing term	63
Another type of initial condition, arising in problems of bore propagation	65
Analyticity of solutions with respect to t	66
4. Uniqueness and stability	67
Uniqueness	68
Stability	68
5. Conclusion	70
Appendix 1. Model equations representing other forms of dispersion	72
Examples	74
Outline of existence theory	74
Appendix 2. Nullity at infinity of a function class	76
Glossary of special spaces	77
References	78

Several topics are studied concerning mathematical models for the unidirectional propagation of long waves in systems that manifest nonlinear and dispersive effects of a particular but common kind. Most of the new material presented relates to the initial-value problem for the equation

$$u_t + u_x + uu_x - u_{xxt} = 0, \qquad (a)$$

whose solution $u(x,t)$ is considered in a class of real nonperiodic functions defined for $-\infty < x < \infty$, $t \geqslant 0$. As an approximation derived for moderately long waves of small but finite amplitude in particular physical systems, this equation has the same formal justification as the Korteweg–de Vries equation

$$u_t + u_x + uu_x + u_{xxx} = 0, \qquad (b)$$

with which (*a*) is to be compared in various ways. It is contended that (*a*) is in important respects the preferable model, obviating certain problematical aspects of (*b*) and generally having more expedient mathematical properties.

The paper divides into two parts where respectively the emphasis is on descriptive and on rigorous mathematics. In § 2 the origins and immediate properties of equations (*a*) and (*b*) are discussed in general terms, and the comparative shortcomings of (*b*) are reviewed. In the remainder of the paper (§§ 3, 4) – which can be read independently of the preceding discussion – an exact theory of (*a*) is developed. In § 3 the existence of classical solutions is proved; and following our main result, theorem 1, several extensions and sidelights are presented. In § 4 solutions are shown

† On leave from the University of Western Australia.

to be unique, to depend continuously on their initial values, and also to depend continuously on forcing functions added to the right-hand side of (a). Thus the initial-value problem is confirmed to be classically well set in the Hadamard sense.

In appendix 1 a generalization of (a) is considered, in which dispersive effects within a wide class are represented by an abstract pseudo-differential operator. The physical origins of such an equation are explained in the style of § 2, two examples are given deriving from definite physical problems, and an existence theory is outlined. In appendix 2 a technical fact used in § 3 is established.

1. Introduction

The equation studied in this paper is useful in that it describes approximately the unidirectional propagation of long waves in certain nonlinear dispersive systems. Such also is the well-known equation of Korteweg & de Vries (1895), concerning which a great deal of new theory has appeared in recent years (see, for instance, Gardner, Greene, Kruskal & Miura 1967; Miura 1968; Miura, Gardner & Kruskal 1968; Lax 1968). Under the assumption of small wave-amplitude and large wavelength, the KdV equation was originally derived for water waves and it is similarly justifiable as a model for long waves in many other physical systems. It has been used to account adequately for observable phenomena such as the interaction of solitary waves and dissipationless, undular shocks. Our main contention in this paper, however, is that the equation under investigation, which stands as a rational alternative to the KdV equation, is in important respects a more satisfactory model.

When the physical parameters and scaling factors presented in a particular example are appropriately absorbed into the definitions of the dependent variable u and the independent variables x and t, which are respectively proportional to distance in the physical system and to time, the KdV equation is obtained in the tidy form

$$u_t + u_x + uu_x + u_{xxx} = 0. \tag{1.1}$$

A further reduction, removing the second term of (1.1), may be made by taking $x' = x - t$ and t as independent variables. Equation (1.1), or its equivalent without the second term, is commonly taken as the starting-point for mathematical studies of long-wave phenomena, although facts with considerable theoretical significance are already entailed in the derivation of (1.1). Thus the condensed form of the KdV equation tends to disguise the meaning of the theory of the equation with regard to the original physical problem. In § 2 we shall review the essentials of the derivation of the KdV equation in particular physical examples, and we shall point out certain theoretical difficulties associated with the equation which arise spuriously, being irrelevant to the original problem. It will be argued in § 2 that in all examples the assumptions leading to the KdV equation equally well justify the equation

$$u_t + u_x + uu_x - u_{xxt} = 0 \tag{1.2}$$

as a model for describing long-wave behaviour, and this alternative obviates the difficulties in question. We shall refer to (1.2) as the regularized long-wave equation, reflecting in this term our view that (1.1) is an unsuitably posed model for long waves.

The preferability of (1.2) over (1.1) finally became clear to us after attempting to formulate an existence theory for (1.1), respective to nonperiodic initial values $u(x, 0)$ defined on the unbounded interval $(-\infty, \infty)$. Notwithstanding the many impressive and seemingly useful properties of solutions of (1.1) that have been demonstrated in recent work, in respect of the central problem of existence the equation poses grave technical difficulties which exclude the possibility of neat results being obtained. Although probably not insuperable, these difficulties appear disproportionate to the status of the equation as an approximate model for long waves. The fact that, in contrast, existence and stability theory for (1.2) is essentially straightforward is perhaps the most persuasive evidence that (1.2) is better founded than (1.1) as a rational model.

For the KdV equation, the question of the existence of *periodic* solutions corresponding to periodic initial values was considered by Sjöberg (1970), but his arguments appear open to question in several respects.† He suggested that his method of analysis might readily be extended to the problem of non-periodic initial values defined on $(-\infty, \infty)$, but a little study shows this suggestion to be unfounded: the type of argument used by him rests crucially on the assumption of a bounded domain of definition. A comparable investigation of the KdV equation has been made by Temam (1969), a full account of whose work may be found in a book by Lions (1969, ch. 3, § 4). Temam proved the existence of periodic solutions by the method of parabolic regularization, first modifying the equation by the addition of a term ϵu_{xxxx} which ensures good properties of solutions, and finally letting $\epsilon \to 0$ in the results for the modified equation. The proof developed by Temam is very intricate and he made no claim that it could be extended to the initial-value problem on the infinite interval. A global existence theorem relating to the problem on the infinite interval has been stated by Kametaka (1969) in a short note, but the detailed proof was deferred to a later publication which we have not yet seen.

In § 3 the existence of non-periodic solutions of the initial-value problem for (1.2) is proved. First, a fixed-point theorem is used to establish existence in the small (i.e. over a sufficiently small interval of time following the initial instant), and then the result is extended to arbitrary time-intervals by appeal to a property that is to be introduced in the next paragraph. In § 4 the uniqueness of solutions is demonstrated, and then stability properties of the regularized equation are studied. It is shown that solutions depend continuously on their initial values, and that the effects of small corrections to the equation remain small over all time-intervals such that this might reasonably be expected. The latter property appears particularly important in that it ensures the validity of the equation as a physical approximation. In appendix 1 a generalization of the problem at issue is reviewed, relating to examples of physical systems in which the dispersion of long waves cannot be modelled by a differential equation. In place of (1.2) the equation in question is

$$u_t + u_x + uu_x + (Hu)_t = 0,$$

where the linear operator H belongs to a class of pseudo-differential operators.

The arguments developed in §§ 3 and 4 depend in large part on the following simple property of solutions $u(x, t)$ of (1.2). We provisionally assume

(i) that $$u, u_x, u_{xt} \to 0 \quad \text{as} \quad x \to \pm\infty,$$

and (ii) that $$E(u) = \int_{-\infty}^{\infty} (u^2 + u_x^2)\, dx \qquad (1.3)$$

† Specifically, his statement of an existence theorem is somewhat indefinite and his proof appears incomplete. Considering the x-interval $[0, l]$ as one period and taking in the first place a suitably small time-interval $[0, t_0]$, he constructs a sequence of periodic functions which is shown to have a limit in the class of functions $u(x, t)$ which are differentiable with respect to t and whose first three derivatives with respect to x belong to $L_2(0, l)$. This limit would be a classical solution of the KdV equation if it had appropriate differentiability properties, but these are not established. In fact Sjöberg does not make clear in his paper what meaning is attached to the term solution. Considering the existence of a solution to have been established for small times, Sjöberg demonstrates existence on an arbitrary time-interval by assuming further differentiability with respect to time and using certain invariant functions of solutions as discussed by Miura *et al.* (1968). The proof of the invariance of these functions seems to depend, however, on the existence of spatial derivatives of higher order than is established by Sjöberg.

exists. These conditions will be confirmed later under appropriate assumptions on the initial data. When (1.2) is multiplied by u and then integrated between $x = -R$ and $x = R$, an integration by parts of the final term on the left-hand side gives

$$\int_{-R}^{R} (uu_t + u_x u_{xt})\,\mathrm{d}x + \left[\tfrac{1}{2}u^2 + \tfrac{1}{3}u^3 - uu_{xt}\right]_{x=-R}^{x=R} = 0. \tag{1.4}$$

Because of (i) the integrated terms vanish in the limit as $R \to \infty$, and hence we have

$$\int_{-\infty}^{\infty} (uu_t + u_x u_{xt})\,\mathrm{d}x = \frac{1}{2}\frac{\mathrm{d}E(u)}{\mathrm{d}t} = 0.$$

Thus

$$E(u) = \text{const.} \tag{1.5}$$

Using a customary designation we shall refer to the functional $E(u)$ as the energy integral, although it is not necessarily identifiable with energy in the original physical problem. Note that the assumption $E(u) < \infty$ means that $u(x, t)$ is for any t an element of the Sobolev space $W_2^1(-\infty, \infty)$, which consists of square-integrable (L_2) functions with generalized first derivatives that are also square-integrable: the norm of this function space is $\|u\| = \sqrt{E(u)}$ (see Smirnov 1964, §§ 112, 114). It is well known that W_2^1 is embedded in the space C of bounded continuous functions (Smirnov, p. 340, theorem 1): that is, if we ignore the equivalence of functions differing from u only on sets of zero measure (as we justifiably can for all present purposes), then $E(u) < \infty$ implies u to be a continuous and bounded function of x on $(-\infty, \infty)$. In fact one can show by a simple argument† that

$$\sup_{-\infty < x < \infty} |u(x, t)| \leqslant \sqrt{E(u)}. \tag{1.6}$$

Another invariant nonlinear functional of solutions of (1.2) is reported in a paper by one of us (Benjamin 1972, appendix B) dealing with the stability of solitary waves.

Other concepts from functional analysis will be introduced where they are required in §§ 3 and 4, and a glossary of special function spaces used in the analysis may be found at the end of the paper. Partial differentiation will mostly be denoted by subscripts, as exemplified in (1.1) and (1.2), but use will also be made of the alternative notation ∂_x, ∂_t when the connotation is thereby made clearer.

2. Discussion of approximate equations for the evolution of long waves

To put the contribution of this paper into focus, we need first to recall the principles of the derivation of the KdV equation in physical examples. This is a matter that has been discussed previously by many writers, including Broer (1964), Benjamin (1967 *a*), Meyer (1967) and Su & Gardner (1969).

The details of the derivation differ, of course, in different examples, but in general the essentials are as follows. We use the notation u^* for the dependent variable as originally presented by the physical problem, likewise x^* and t^* for physical distance and time. The first essential property of the systems in question is that dispersive effects on *infinitesimal* waves vanish in the limit as wavelength becomes infinite, and the limiting phase speed is a constant $c_0 > 0$. Thus, respecting

† Considering the integral of the non-negative function $(u-u_x)^2$. By considering the Fourier transforms of the L_2 functions u and u_x, and using Parseval's theorem for integrals, one may derive an estimate sharper than (1.6) in that a factor $2^{-\frac{1}{2}}$ multiplies the right-hand side (cf. Benjamin 1972, § 3). But (1.6) suffices for present purposes, and is preferred for the sake of tidiness.

infinitesimal waves of extreme length, propagation in the $+x^*$ direction is described by the equation

$$u^*_{t^*} + c_0 u^*_{x^*} = 0, \tag{2.1}$$

whose general solution is an arbitrary differentiable function of $x^* - c_0 t^*$. It deserves emphasis that the assumption of *unidirectional* propagation, introduced with obvious meaning here, will be no less essential to the modifications of (2.1) presently in question. (We should note that the unidirectional nature of the present model equations is generally not essential to the original physical systems which these equations are meant to simulate approximately, so that special justification may be needed for this attribute of the modelling. Such justification is to be found only in the context of particular physical problems, however, and we shall not go into this aspect here.) In certain applications to nonlinear dispersive systems (e.g. to water waves of tidal proportions), equation (2.1) already has some validity as an approximation for real waves of sufficiently small amplitude and great length; but it is not a valid approximation over very large times, during which nonlinear and frequency-dispersive effects can accumulate to a significant level. Accordingly, the object of first-order theories improving on this rudimentary approximation is to establish corrections to (2.1) which represent with a reasonably extended range of validity, first, nonlinear effects on waves of finite but small amplitude and, secondly, dispersive effects as suffered by waves of finite but large wavelength. Although for a self-consistent theory allowance has to be made for the two kinds of small effect simultaneously, the outcome of such theories can generally be anticipated by considering the two effects separately. That is, in a first approximation each is separately accountable by a small correction added to (2.1), and higher-order error terms whose estimation would entail consideration of interactions between the two appear negligible in this approximation.

If the effects of finite wavelength are ignored, it is found, as another general attribute of the systems in question, that small nonlinear effects on waves propagating in the $+x^*$ direction are representable approximately in the following way. Whereas according to the linearized equation (2.1) all specific values of u^* are propagated along characteristics with the same velocity $\mathrm{d}x^*/\mathrm{d}t^* = c_0$, at the next approximation the characteristic velocity becomes dependent linearly on u^*: thus

$$\frac{1}{c_0}\left(\frac{\mathrm{d}x^*}{\mathrm{d}t^*}\right)_{u^* = \text{const.}} = 1 + bu^*, \tag{2.2}$$

where b is a constant. This property generally depends on $|bu^*|$ being small, so we clarify its meaning by writing

$$bu^* = \epsilon U \quad (\epsilon > 0), \tag{2.3}$$

and considering the scaled dependent variable U to be of unit order of magnitude. The validity of (2.2) then rests on the conditions that the amplitude parameter ϵ is sufficiently small ($\epsilon \ll 1$) and that the implicit error is $o(\epsilon)$. In fact the error is generally found to be $O(\epsilon^2)$. After this substitution, (2.2) is precisely equivalent to

$$U_{t^*} + c_0 U_{x^*} + \epsilon c_0 U U_{x^*} = 0, \tag{2.4}$$

which we can regard as the improvement on (2.1) that accounts approximately, to $O(\epsilon)$, for nonlinear effects (cf. Lamb 1932, § 187).

If, on the other hand, the effects of finite amplitude are ignored, the (linearized) theory of travelling waves can generally be developed without restriction on wavelength. In respect of simple-harmonic waves represented by $u^* = a \exp(\mathrm{i}\sigma t - \mathrm{i}\kappa x)$, the properties of the physical system are found to determine a *dispersion relation* between frequency σ and wave number κ, and this may be obtained explicitly, or at least implied, in the form

$$\sigma = \kappa c(\kappa), \tag{2.5}$$

where the phase velocity c is expressed as a function of κ. The function $c(\kappa)$ is necessarily even and is generally positive for all κ, so that only waves travelling in the $+x^*$ direction are represented as above. As already explained, we have $c(0) = c_0$. The result for simple-harmonic waves may be generalized by Fourier's principle, on the assumption that the solution $u^*(x^*, t^*)$ of the linearized problem and its derivatives are, for each t^*, functions of x^* on $(-\infty, \infty)$ to which the Fourier integral theorem is applicable. Thus one infers that u^* satisfies the equation

$$u^*_{t^*} + c_0(Lu^*)_{x^*} = 0, \tag{2.6}$$

in which L is the linear transformation defined by

$$Lu^* = \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{c(\kappa)}{c_0}\, e^{i\kappa(x^*-\xi)} u^*(\xi, t^*)\, d\xi\, d\kappa. \tag{2.7}$$

The next step is to establish an approximation to L valid for long waves.

For the systems in question, the essential property used at this stage of the argument is that $c(\kappa)$ has a smooth maximum with non-vanishing curvature at $\kappa = 0$.† That is, an approximation for sufficiently small κ is

$$c(\kappa) = c_0(1 - \alpha^2\kappa^2), \tag{2.8}$$

where

$$c_0\alpha^2 = -\tfrac{1}{2}c''(0) > 0.$$

This suggests the introduction of the scaled independent variables

$$X = \epsilon^{\frac{1}{2}}x^*, \quad T = \epsilon^{\frac{1}{2}}c_0 t^*, \tag{2.9}$$

and the substitution $u^* = (\epsilon/b)\, U(X, T)$ which transforms (2.6) into

$$U_T + (L_\epsilon U)_X = 0, \tag{2.10}$$

with

$$L_\epsilon U = \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{c(\epsilon^{\frac{1}{2}}K)}{c_0}\, e^{iK(X-\Xi)} U(\Xi, T)\, d\Xi\, dK. \tag{2.11}$$

The simplest argument here, which is essentially the one that has most commonly been used in specific examples, supposes that all X-derivatives of $U(X, T)$ are of the same (unit) order of magnitude as U itself. It also has to be assumed in the present approach that these derivatives are square-integrable on $(-\infty, \infty)$, so having Fourier transforms. Then, if the even function $c(\epsilon^{\frac{1}{2}}K)/c_0$ is expandable as a Maclaurin series

$$\frac{c(\epsilon^{\frac{1}{2}}K)}{c_0} = 1 + \sum_{n=1}^{\infty} A_n \epsilon^n K^{2n}, \quad (A_1 = -\alpha^2), \tag{2.12}$$

which is absolutely convergent for all $\epsilon^{\frac{1}{2}}K$, the expression (2.11) is formally equivalent to

$$L_\epsilon U = U + \sum_{n=1}^{\infty} (-1)^n A_n \epsilon^n \partial_X^{2n} U, \tag{2.13}$$

which for $\epsilon \ll 1$ will be a rapidly convergent series by virtue of the absolute convergence of the series (2.12) and the assumption about the magnitudes of the derivatives of U. Truncating (2.13) at first order in ϵ, we have the approximation

$$L_\epsilon U = U + \epsilon\alpha^2 U_{XX}; \tag{2.14}$$

† For example, the form of $c(\kappa)$ applicable to surface waves on water of depth h is

$$c(\kappa) = c_0\left(\frac{\tanh \kappa h}{\kappa h}\right)^{\frac{1}{2}} = c_0\{1 - \tfrac{1}{6}(\kappa h)^2 + \tfrac{19}{360}(\kappa h)^4 \mp \dots\},$$

where $c_0 = (gh)^{\frac{1}{2}}$.

and so (2.9) becomes

$$U_T + U_X + \epsilon\alpha^2 U_{XXX} = 0, \tag{2.15}$$

which we can regard as the improvement on (2.1) accounting for dispersive effects to first order in ϵ.

The foregoing explanation of the approximate dispersion equation (2.15) perhaps suffices to indicate essentials. It is worth noting, however, how (2.15) can be justified more precisely and at the same time the assumptions about the long-wave solutions U of (2.10) can be weakened considerably. Considering the remainder left after the approximation (2.14), that is,

$$\Delta U = L_\epsilon U - U - \epsilon\alpha^2 U_{XX},$$

we wish to establish that $(\Delta U)_X$ is $O(\epsilon^2)$, so presumably being insignificant compared with the terms in (2.15) if ϵ is sufficiently small. The definition (2.11) implies that

$$\Delta U = \frac{1}{2\pi}\int_{-\infty}^{\infty} \gamma(\epsilon^{\frac{1}{2}}K)\, \mathrm{e}^{\mathrm{i}KX}\, \hat{U}(K, T)\, \mathrm{d}K,$$

where $\hat{U}(K, T)$ is the Fourier transform of $U(X, T)$ in X and

$$\gamma(\kappa) = \frac{c(\kappa)}{c_0} - 1 + \alpha^2\kappa^2.$$

Hence, by use of the inequality (1.6) combined with Parseval's theorem, it follows that

$$\sup_{-\infty<X<\infty} |(\Delta U)_X| \leqslant \left\{\frac{1}{2\pi}\int_{-\infty}^{\infty} (1+K^2)\, K^2 |\widehat{\Delta U}|^2\, \mathrm{d}K\right\}^{\frac{1}{2}}$$

$$= \left\{\frac{1}{2\pi}\int_{-\infty}^{\infty} \gamma^2(\epsilon^{\frac{1}{2}}K)\, (K^2+K^4)\, |\hat{U}|^2\, \mathrm{d}K\right\}^{\frac{1}{2}}.$$

In real examples $c(\kappa)$ is a positive function decreasing monotonically with increasing $|\kappa|$, and accordingly a finite constant B can be found such that $|\gamma(\kappa)| \leqslant B\kappa^4$ for all κ. Evidently the least possible value for such a constant is $B = (1/4!)\, c^{\mathrm{iv}}(0)$, but in several examples the authors have tested this value suffices. Thus we conclude that

$$\sup_{-\infty<X<\infty} |(\Delta U)_X| \leqslant \epsilon^2 B \left\{\frac{1}{2\pi}\int_{-\infty}^{\infty} (K^{10}+K^{12})\, |\hat{U}|^2\, \mathrm{d}K^{\frac{1}{2}}\right\}^{\frac{1}{2}}$$

$$= \epsilon^2 B \left\{\frac{1}{2\pi}\int_{-\infty}^{\infty} (U_{5X}^2 + U_{6X}^2)\, \mathrm{d}X\right\}^{\frac{1}{2}}.$$

A sufficient condition for (2.14) to be a valid approximation for small ϵ is, therefore, that both the fifth and sixth derivatives, U_{5X} and U_{6X}, are square integrable. This implies that U_{5X} is bounded, which is obviously a necessary condition.

(It should be acknowledged at this point that in other examples of physical systems the dispersion relation does not admit a small-κ approximation in the form (2.8), so that the present conclusions do not apply. The general case will be discussed in appendix 1.)

Having obtained (2.4) and (2.15) as the respective first-order approximations by allowing separately for small nonlinear and dispersive effects, we may plausibly argue that an approximation accounting simultaneously for both factors is given simply by combining the ϵ terms from (2.4) and (2.14). Thus, changing to the scaled variables X, T in (2.4), we anticipate the equation

$$U_T + U_X + \epsilon(UU_X + \alpha^2 U_{XXX}) = 0, \tag{2.16}$$

which has been obtained in various examples where a unified approximation for nonlinear and dispersive effects has been developed. The parameter ϵ having served its purpose, the approximate equation (2.16) may be reduced by the introduction of the dimensionless but *unscaled* variables

$$\left.\begin{aligned} x = \epsilon^{-\frac{1}{2}}X/\alpha = x^*/\alpha, \quad t = \epsilon^{-\frac{1}{2}}T/\alpha = c_0 t^*/\alpha, \\ u = \epsilon U = bu^*. \end{aligned}\right\} \tag{2.17}$$

This gives us the tidy form (1.1) of the KdV equation; but in thus dispensing with ϵ we recognize a need for caution, because the essential conditions on the magnitudes of the dependent variable and its derivatives are now hidden.

A point of central importance to our discussion is that, to the first order in ϵ, (2.15) is equivalent to

$$U_T + U_X + \epsilon(UU_X - \alpha^2 U_{XXT}) = 0, \tag{2.18}$$

since $U_X = -U_T$ to zero order. Hence the same reduction, in terms of the variables (2.17), leads to the alternative equation introduced in § 1, namely

$$u_t + u_x + uu_x - u_{xxt} = 0. \tag{2.19}$$

It is worth further emphasis that, as an approximate model for long waves of small amplitude, (2.19) has essentially the same formal justification as the KdV equation.† Significant differences appear, however, with regard to basic mathematical and computational aspects of the two equations, so that their effectiveness as models can in several respects be markedly contrasted.

To recall the kind of question generally arising in the assessment of such mathematical models, consider the requirement that an initial-value problem modelling a physical situation should be well set according to Hadamard's definition (1923). That is, for some general class of initial data, a unique 'classical' (i.e. appropriately differentiable) solution should exist which depends continuously on the initial data. Another such practically important requirement is an adequately extended range of validity: the error implicit in the model, when represented by a small term added to the approximate equation, should have for a reasonably long time only a small effect on the solution. A further, obviously desirable element of good mathematical modelling is that the latter properties should be demonstrable without undue difficulty, and it is particularly in this respect, we contend, that the regularized long-wave equation (2.19) is preferable to the KdV equation. A satisfactory theory for the KdV equation, establishing the desirable properties mentioned above, is not yet available, and it would certainly present a much more difficult task of analysis than the theory for (2.19) that is developed in the subsequent sections of this paper. We continue the discussion here by examining several problematical aspects of the KdV equation, which are seen to be obviated by the regularized equation.

Shortcomings of the KdV equation

It is not intended to discuss in any detail the problems of computing solutions, but in passing a few comments on this aspect are deserved. Numerical solutions of the regularized equation (2.19) have been presented by Peregrine (1964), who was careful to choose initial data compatible with the assumptions of the model, so ensuring that his solutions remained physically relevant. His finite-difference scheme gave little trouble. On the other hand, many numerical studies of the KdV equation (1.1) have been reported, but most of this work is open to criticism on the grounds that the values of the solution u and its derivatives were allowed to be unduly large. A common

† In the problem of long water waves, an approximate equation governing unidirectional propagation is commonly derived by way of the so-called Boussinesq equations, a pair of simultaneous partial differential equations in which the dependent variables are the vertical displacement of the free surface and the mean horizontal velocity of the water. In this approach, equation (2.19) in fact comes out more 'naturally' than the formally equivalent KdV equation (cf. Peregrine 1964).

source of difficulty seems to be that, for von Neumann stability of the finite-difference scheme, the spatial step-length Δx and time step Δt need to be chosen so that $(\Delta x)^3/\Delta t$ is a fairly large number, say about 20,† but in addition the values of $|u|\,\Delta t/\Delta x$ have not to be too small. Consequently rather large values of $|u|$ (sometimes considerably greater than 1) have had to be allowed. The physical relevance of the results is then questionable, of course, in that the validity of the KdV equation generally depends on the condition $|u| \ll 1$.

Much recent work on the KdV equation has been concerned with associated conservation laws in the form

$$T_t + X_x = 0, \tag{2.20}$$

where the 'density' T and 'flux' X are polynomials in the solution u and its x-derivatives. The KdV equation itself has this form (i.e. $T_0 = u$, $X_0 = u + \frac{1}{2}u^2 + u_{xx}$), and progressively more complicated pairs T_n, X_n have been given explicitly from $n = 1$ to $n = 9$ by Miura *et al.* (1968). The assumption that $X_n \to 0$ as $x \to \pm\infty$ implies that

$$\int_{-\infty}^{\infty} T_n \, \mathrm{d}x = \text{const.}, \tag{2.21}$$

and thus one has a functional that for solutions of the KdV equation is invariant with time. It has been shown by Miura *et al.* that, for C^∞ solutions which vanish together with all their x-derivatives as $x \to \pm\infty$, there exist infinitely many such invariants, and this remarkable fact appears at first sight very favourable to the prospects of regularity theory for the KdV equation. As regards basic questions of the existence of solutions corresponding to general classes of initial values, the outstanding potentiality of invariance properties like (2.21) is that they may provide useful *a priori* estimates, such as may enable local (small-time) results to be extended to global results. The apparently favourable situation in having any number of these properties available is somewhat deceptive, however, as the following simple example shows.

Consider the conservation law in the form (2.20) obtained by multiplication of (1.1) by u: that is,

$$T_1 = \tfrac{1}{2}u^2,$$

$$X_1 = \tfrac{1}{2}u^2 + \tfrac{1}{3}u^3 + uu_{xx} - \tfrac{1}{2}u_x^2.$$

If the derivative $\partial_x X_1$ exists (at least in a generalized L_2 sense), so that T_1 is correspondingly differentiable with respect to t, and if X_1 vanishes as $x \to \pm\infty$, then we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{-\infty}^{\infty} \tfrac{1}{2}u^2 \, \mathrm{d}x = 0. \tag{2.22}$$

Thus the L_2 norm of a solution keeps its initial value. But there are, of course, L_2 functions for which X_1 is not differentiable in even the L_2 sense, and for which none of the four components of X_1 vanishes at infinity. The invariant (1.5) for solutions of the regularized equation presents a similar situation in that it depends on certain conditions at infinity being satisfied, but these are readily verifiable by a local existence theory (see § 3). In contrast, the conditions justifying (2.22) appear very difficult to establish as a consequence of the initial data alone. We note that they might be established by showing that higher-order functionals in the form (2.21) remained bounded; however, at least the next three (in the hierarchy defined by Miura *et al.*) would need to be so restricted, and the invariance of these depends in turn on further differentiability requirements and conditions at infinity.

† This serves to obliterate spurious short-wave behaviour as discussed below.

It is helpful to recognize that the main difficulties presented by the KdV equation arise from the dispersion term and so are common to the linearized equation

$$u_t + u_x + u_{xxx} = 0, \tag{2.23}$$

which can more readily be used to illustrate their nature. First note that when the solution of (2.23) is expressible as a summation of Fourier components in the form $F(k)\,e^{i\omega t - ikx}$, the dispersion relation determined by (2.23) is

$$\omega = k - k^3,$$

which corresponds, of course, to the approximation (2.8) justified only for small wavenumbers. Unless the spectrum $F(k)$ is cut off beyond small values of k, the solution therefore contains essentially artificial features, and it is the unduly potent behaviour of these that makes the model troublesome. The phase velocity ω/k becomes negative for $k^2 > 1$, in contradiction of the original assumption of forward-travelling waves. More significantly, the group velocity

$$d\omega/dk = 1 - 3k^2$$

has no lower bound, which means that no limit can be assigned to the rate at which fine-scale features of the solution are transmitted in the $-x$ direction. This can be troublesome computationally even when the initial waveform is genuinely long; for round-off errors will generally add small-scale features with which the model will then have to contend.

Let us consider the solution of (2.23) corresponding to given initial data on the whole real axis x: that is, we require $u(x,t)$ to satisfy (2.23) for $t > 0$, $-\infty < x < \infty$, and to have the property $u \to g(x)$ as $t \downarrow 0$, where $g(x)$ is a given function. It may readily be confirmed that the solution is

$$u = \frac{1}{(3t)^{\frac{1}{3}}} \int_{-\infty}^{\infty} \mathrm{Ai}\left(\frac{x-t-s}{3^{\frac{1}{3}} t^{\frac{1}{3}}}\right) g(s)\, ds, \tag{2.24}$$

provided this integral exists. Here Ai denotes the Airy function, normalized so that

$$\int_{-\infty}^{\infty} \mathrm{Ai}(z)\, dz = 1.$$

In keeping with the property of the group velocity noted above, the Green function appearing in (2.24) has fiercely oscillatory behaviour for large negative arguments: specifically, the asymptotic form in question is $\mathrm{Ai}(-z) \sim \pi^{-\frac{1}{2}} z^{-\frac{1}{4}} \sin(\frac{2}{3} z^{\frac{3}{2}} + \frac{1}{4}\pi)$. Consequently, considered as a transformation of the initial waveform $g(x)$, the expression (2.24) has a lack of continuity and tendency to emphasize short-wave components that is unnatural in respect of the original physical problem. In the literature on long waves, various example solutions of (2.23) have been given which incidentally feature short-wave components, but for which, accordingly, special considerations have to be made (cf. Jeffreys & Jeffreys 1966, § 17.09; Benjamin & Barnard 1964, § 4). In fact, to impart some degree of realism to such examples, it seems essential that short-wave features of the initial waveform should be put at or behind its main front, not far ahead. The short waves are then soon transmitted far to the rear of the evolving wave-front, which becomes the part of the solution that is meaningful as a physical approximation. In general, however, short waves introduced in the initial waveform may as they evolve become concentrated in such a way as to make the solution totally unacceptable. For example, consider

$$g(x) = \frac{\mathrm{Ai}(-\beta x)}{(1+x^2)^m},$$

with $\beta > 0$ and $\frac{1}{8} < m < \frac{1}{4}$. This is a continuous function which is also an element of $L_2(-\infty, \infty)$. But one finds that at the point $x = t$ the solution (2.24) does not exist in the limit as $t \to 1/(3\beta^3)$.

We now observe that the linearized version of the regularized equation (2.19) presents none of these awkward properties. The dispersion relation determined by it is seen to be

$$\omega = \frac{k}{1+k^2}, \tag{2.25}$$

according to which both the phase velocity ω/k and group velocity $d\omega/dk$ are bounded for all k. Moreover, both velocities approach zero for large k, which implies that fine-scale features of the solution tend not to propagate. In other words, the model has the desirable property of responding only feebly to artificial short-wave components that may be introduced in the initial waveform (or in a computational procedure). It will be verified precisely in § 3 that the same property is possessed by the nonlinear model, equation (2.19).

A comparison deserves to be made at this point with a type of model proposed by Whitham (1967), in which the exact (linear) dispersive properties of a system are represented but nonlinear effects are simulated only by the first-order approximation appropriate to long waves. Thus, after introducing dimensionless variables, one obtains the equation

$$u_t + (Lu)_x + uu_x = 0, \tag{2.26}$$

where L is the linear operator defined in general by (2.7). Being exact, the linearized form of this equation has properties like those just mentioned, which were pointed out to arise from the artificial but well-behaved dispersion relation (2.25). Unlike (2.19), however, the nonlinear model (2.26) does not preserve such good short-wave properties. Whereas dispersive effects disappear for very short waves, the artificial nonlinear effects do not; and so for these waves there is a tendency towards shock formation – in the same manner as is familiar for extremely long waves – which is unresisted by dispersion. It must be expected that for any initial waveform with significant short-wave components, the solution of (2.26) will soon cease to be single-valued.

To compare (2.19), it is appropriate to introduce a new dependent variable defined by the relation

$$v = u - u_{xx} \quad (v(\pm\infty, t) = u(\pm\infty, t) = 0),$$

the inverse of which is

$$u = \int_{-\infty}^{\infty} \tfrac{1}{2} e^{-|x-s|} v(s,t)\, ds$$
$$= Kv, \quad \text{say.}$$

Equation (2.19) then takes the form

$$v_t + (Kv)_x + (Kv)(Kv)_x = 0. \tag{2.27}$$

The linear operator K is, of course, just that corresponding to the dispersion relation (2.25), which we saw to indicate that dispersive effects on short-wave components become progressively feebler with their fineness of scale. We now see from (2.27) that nonlinear effects are correspondingly enfeebled, so that the unsatisfactory short-wave behaviour of the model (2.26) does not arise.

3. Existence theory for the regularized equation

Directing our attention now to the model equation that was introduced as (1.2) and (2.19), we shall investigate the initial-value problem for it on the unbounded domain $\mathbb{R}$ (i.e. $-\infty < x < \infty$). The existence of a classical solution† will be established upon the assumptions that the initial

† By this term we mean a solution that is continuously differentiable to the orders required by the partial differential equation, and satisfies the equation pointwise in a given domain of x and t.

waveform $u(x, 0)$ is smooth enough for the partial differential equation to be meaningful, that this waveform is evanescent in a certain way at infinity, and that the initial 'energy' – as measured by the functional (1.3) – is finite. Following a common procedure, we shall first prove a local existence theorem (i.e. for a sufficiently small time-interval) by means of a fixed-point principle; and properties of the solution thus established will then be shown to substantiate an energy estimate which provides a starting-point for a further application of the local theorem. It will be seen that the argument can be repeated indefinitely, so establishing the existence of a solution over an arbitrary time-interval. After the presentation of our main result (theorem 1), several sidelights and extensions will be discussed.

To provide an amenable version of the problem, the differential equation (1.2) first needs to be converted by formal operations into an integral equation. Rewriting the equation as

$$(1-\partial_x^2)\, u_t = -\partial_x(u+\tfrac{1}{2}u^2),$$

we may regard it as an ordinary differential equation for u_t, the formal solution of which is

$$u_t = -\frac{1}{2}\int_{-\infty}^{\infty} \mathrm{e}^{-|x-\xi|}\,\partial_\xi\{u(\xi,t)+\tfrac{1}{2}u^2(\xi,t)\}\,\mathrm{d}\xi.$$

After a formal integration by parts, this becomes

$$u_t = \int_{-\infty}^{\infty} K(x-\xi)\,(u+\tfrac{1}{2}u^2)\,\mathrm{d}\xi,$$

with

$$K(x) = \tfrac{1}{2}(\operatorname{sgn} x)\,\mathrm{e}^{-|x|},$$

from which there follows

$$u(x,t) = g(x) + \int_0^t \int_{-\infty}^{\infty} K(x-\xi)\,\{u(\xi,\tau)+\tfrac{1}{2}u^2(\xi,\tau)\}\,\mathrm{d}\xi\,\mathrm{d}\tau, \tag{3.1}$$

where

$$g(x) = u(x,0).$$

We write this integral equation for short as

$$u = Au = g + Bu. \tag{3.1'}$$

Our aim, achieved through a sequence of lemmas, is to show subject to suitable restrictions on the function $g(x)$ that (3.1) has a solution which is simultaneously a solution of the original differential equation.

Notation. We write $\mathscr{C}_T[\equiv C(\mathbb{R}\times[0,T])]$ to denote the class of functions $v(x,t)$ that are continuous and uniformly bounded on the infinite strip $\mathbb{R}\times[0,T]$ (i.e. on the set of points (x,t) such that $-\infty < x < \infty$, $0 \leqslant t \leqslant T$). An instance of this function class is considered in the proof of the following lemma as a Banach space. We also write $\mathscr{C}_T^{l,m}$ for the narrower class of functions $v(x,t)$ such that $\partial_x^i\partial_t^j v \in \mathscr{C}_T$ for $0 \leqslant i \leqslant l$, $0 \leqslant j \leqslant m$. These and other special symbols are listed in the glossary at the end of the paper. When $\mathscr{C}$ appears as a subscript to the norm symbol $\|\,.\,\|$, indicating that the norm is for the space $\mathscr{C}_T$, the further subscript T is omitted. In all cases it will be clear from the context what t-interval is implied.

Lemma 1. Let $g(x)$ be a continuous function such that

$$\sup_{x\in\mathbb{R}} |g(x)| \leqslant b < \infty. \tag{3.2}$$

Then there exists a $t_0(b) > 0$ such that the integral equation (3.1) has a solution, satisfying $u(x,0) = g(x)$, which is bounded and continuous for $x\in\mathbb{R}$ and $0 \leqslant t \leqslant t_0$.

Proof. Consider the Banach space $\mathscr{C}_{t_0}$ of continuous and bounded functions defined on $\mathbb{R}\times[0, t_0]$, with the norm

$$\|v\|_{\mathscr{C}} = \sup_{\substack{x\in\mathbb{R}\\ 0\leqslant t\leqslant t_0}} |v(x,t)|.$$

For the moment, the positive number t_0 entailed in the definition of $\mathscr{C}_{t_0}$ is left arbitrary. We first notice that, for any $v_1, v_2 \in \mathscr{C}_{t_0}$ and any $x\in\mathbb{R}$, $t\in[0,t_0]$,

$$\begin{aligned}|Av_1 - Av_2| = |Bv_1 - Bv_2| &\leqslant \|v_1-v_2\|_{\mathscr{C}}\{1+\tfrac{1}{2}\|v_1+v_2\|_{\mathscr{C}}\}\int_0^t\int_{-\infty}^{\infty} |K(x-\xi)|\,d\xi\,d\tau \\ &= \|v_1-v_2\|_{\mathscr{C}}\{1+\tfrac{1}{2}\|v_1+v_2\|_{\mathscr{C}}\}\,t \\ &\leqslant \|v_1-v_2\|_{\mathscr{C}}\{1+\tfrac{1}{2}\|v_1\|_{\mathscr{C}}+\tfrac{1}{2}\|v_2\|_{\mathscr{C}}\}\,t.\end{aligned}$$

(Here the first inequality is evident from the factorization $v_1^2 - v_2^2 = (v_1-v_2)(v_1+v_2)$, and finally the triangle inequality for norms is used.) Hence, taking the supremum of both sides for $x\in\mathbb{R}$ and $t\in[0,t_0]$, we obtain

$$\|Av_1 - Av_2\|_{\mathscr{C}} \leqslant t_0(1+\tfrac{1}{2}\|v_1\|_{\mathscr{C}}+\tfrac{1}{2}\|v_2\|_{\mathscr{C}})\,\|v_1-v_2\|_{\mathscr{C}}, \tag{3.3}$$

from which it can be confirmed that the operator A is a *continuous* mapping of the space $\mathscr{C}_{t_0}$ into itself. Moreover, it follows that the mapping of the ball $\|v\|_{\mathscr{C}} \leqslant R$ satisfies a Lipschitz condition with Lipschitz constant $\theta < 1$ if

$$t_0(1+R) \leqslant \theta < 1. \tag{3.4}$$

This condition on t_0 and R also implies that $\|Bv\|_{\mathscr{C}} \leqslant \theta\|v\|_{\mathscr{C}}$, as appears when we put $v_2 \equiv 0$ in (3.3); and from this combined with the condition (3.2) of the lemma it is seen that the ball is mapped into itself if, in addition to (3.4),

$$b \leqslant (1-\theta)R. \tag{3.5}$$

The conditions (3.4) and (3.5) together mean that the mapping of the ball is *contractive*. Therefore, according to the principle of contractive mappings (Smirnov 1964, § 86), these conditions ensure that A has a unique fixed point u in the ball $\|v\|_{\mathscr{C}} \leqslant R$.

It merely remains to check that θ, R and t_0 can be chosen to satisfy (3.4) and (3.5) simultaneously. For instance, take $\theta = \frac{1}{2}$ and $R = 2b$, so that (3.5) is satisfied. Then any positive value $t_0 < 1/(2+4b)$ is seen to satisfy (3.4). Thus, as the fixed point of A is a solution of equation (3.1), the lemma is proven. An implication of this proof is that the solution u is obtainable by successive approximations starting from any element of the ball $\|v\|_{\mathscr{C}} \leqslant R$. In particular, we note that the Picard sequence $\{v_n\}$ generated by the formula $v_n = Av_{n-1}$, $v_1 = g$ is convergent in $\mathscr{C}_{t_0}$ towards the limit u.

LEMMA 2. If $g\in C^2(\mathbb{R})$,† then any solution $u(x,t)$ of (3.1) which is an element of $\mathscr{C}_T$ (for a given $T>0$) is also an element of $\mathscr{C}_T^{2,\infty}$ and accordingly is a classical solution of the initial value problem for the partial differential equation (1.2).

Proof. We use typical 'bootstrap' arguments, exploiting the fact that u is identical with Au. Under the conditions of the lemma, Au is a continuously differentiable function of t. Therefore u_t exists, being given by

$$u_t = (Au)_t = \int_{-\infty}^{\infty} K(x-\xi)\{u(\xi,t)+\tfrac{1}{2}u^2(\xi,t)\}\,d\xi, \tag{3.6}$$

† This connotes that the functions $g(x)$, $g'(x)$, $g''(x)$ are each bounded as well as continuous on the infinite interval $\mathbb{R}$.

and is thus continuous in x, t and bounded on $\mathbb{R}\times[0, T]$. Hence it can be argued inductively that the mth derivative ($m = 2, 3, \ldots$) with respect to t also exists, being given by

$$\partial_t^m u = \int_{-\infty}^{\infty} K(x-\xi)\, \partial_t^{m-1}(u+\tfrac{1}{2}u^2)\, d\xi, \tag{3.7}$$

and is continuous and bounded on $\mathbb{R}\times[0, T]$. (Here we use the obvious fact that u^2 has the same degree of regularity as u if u is bounded.) Since clearly the induction can be carried to any m, the statement of the lemma concerning the t-dependence of u is verified.

Next, by dividing the range of integration in (3.1) at $\xi = x$, we confirm the existence of u_x, which is given by

$$u_x = g'(x) + \int_0^t (u+\tfrac{1}{2}u^2)\, d\tau - \int_0^t \int_{-\infty}^{\infty} \tfrac{1}{2}\, e^{-|x-\xi|}(u+\tfrac{1}{2}u^2)\, d\xi\, d\tau. \tag{3.8}$$

This shows that u_x is continuous and bounded, which in turn implies that the first integral on the right-hand side of (3.8) is a continuously differentiable function of x – as obviously is the second integral. So u_{xx} exists, being given by

$$\begin{aligned} u_{xx} &= g''(x) + \int_0^t (u+\tfrac{1}{2}u^2)_x\, d\tau + \int_0^t\int_{-\infty}^{\infty} K(x-\xi)\,(u+\tfrac{1}{2}u^2)\, d\xi\, d\tau \\ &= g''(x) + \int_0^t (u+\tfrac{1}{2}u^2)_x\, d\tau + u - g(x), \end{aligned} \tag{3.9}$$

and is also continuous and bounded. For all the t-derivatives of u_x and u_{xx}, continuity and boundedness are deducible inductively as above. It follows that the solution $u(x, t)$ of (3.1) satisfies (1.2) pointwise in $\mathbb{R}\times[0, T]$, and thus the proof of the lemma is complete.

Under the assumptions of the lemma the foregoing argument cannot be extended to establish the existence of higher x-derivatives of u; for the solution evidently cannot acquire, as a function of x, better regularity properties than those of the initial waveform $g(x)$. It is easy to see, however, that further differentiability conditions on $g(x)$ will correspondingly restrict the solution. Specifically, if $g\in C^l(\mathbb{R})$ with $l > 2$, then continuation of the bootstrap argument shows that $u\in\mathscr{C}_T^{l,\infty}$; and this conclusion still holds if $l = \infty$. Note, incidentally, that if g is only *piecewise* twice-differentiable, our argument still proves that $u - g = Bu$ belongs to $C^2(\mathbb{R})$ for every $t\in[0, T]$. Thus discontinuities introduced initially in u_{xx} *do not propagate*. In fact they remain undiminished in strength at their original positions, while the C^2 part of the solution evolves away from them. Thus, for example, if u_{xx} is discontinuous at $x = 0$, then $u_{xx}(0^+, t) - u_{xx}(0^-, t)$ is independent of t, even though $u(0, t)$, $u_x(0, t)$ and $u_{xx}(0^+, t)$ may all vary.

Our next object is to establish the invariant property (1.5) for the local solution guaranteed by lemma 1. For this purpose we need an unambiguous criterion whereby functions may be specified to vanish at infinity. Accordingly, we shall say that a function $v(x)$ defined on the whole of $\mathbb{R}$ is *asymptotically null* if $\lim_{x\to\infty}|v(x)|$ and $\lim_{x\to-\infty}|v(x)|$ both exist and are equal to zero. The following two properties of asymptotically null functions will be used:

(i) If $\{v_n\}$ is a sequence of functions converging in the Banach space $\mathscr{C}_{t_0}$, as considered in the proof of lemma 1, and if each function is asymptotically null, then so is the limit u of the sequence. This result appears immediately on consideration of the inequality

$$|u| \leqslant |v_n| + |u - v_n|,$$

which obviously must be satisfied for all $x\in\mathbb{R}$ and all n. By choice of n large enough, the l.u.b. of $|u-v_n|$ can be made arbitrarily small, and $|v_n|$ can be made arbitrarily small by choice of $|x|$ large enough. It follows that u must be asymptotically null.

(ii) If the function v is continuous and asymptotically null, then so are both

$$\int_{-\infty}^{\infty} e^{-|x-\xi|} v(\xi)\, d\xi \quad \text{and} \quad \int_{-\infty}^{\infty} K(x-\xi)\, v(\xi)\, d\xi.$$

The continuity of these two functions is obvious, having already been considered in the proofs of lemmas 1 and 2, and the asymptotic property of the second appears directly from the following argument regarding the first. To prove the required property for $x \to \infty$, we may divide the range of integration, say at $\xi = \xi_1$. For $x \geqslant \xi_1$, we then have the estimate

$$\int_{-\infty}^{\infty} e^{-|x-\xi|} v(\xi)\, d\xi \leqslant e^{-x} \int_{-\infty}^{\xi_1} e^{\xi} |v(\xi)|\, d\xi + 2 \sup_{\xi \geqslant \xi_1} |v(\xi)|.$$

The second term on the right-hand side of this inequality can be made arbitrarily small by choosing a sufficiently large finite value of ξ_1, after which the first term can be made arbitrarily small by taking x large enough. The complementary property for $x \to -\infty$ may be demonstrated in the same way.

We can now establish the following proposition:

LEMMA 3. If $u(x,t)$ is the solution of (3.1) assured by lemma 1, and if $g, g', \ldots, g^{(p)}$ are continuous and asymptotically null, then $\partial_x^l \partial_t^m u$ is asymptotically null for all $m \geqslant 0$ and $0 \leqslant l \leqslant p$.

The proof is immediate in the light of the aforesaid properties (i) and (ii), once it is recognized that the solution of the integral equation is the limit of the Picard sequence given by $v_n = Av_{n-1}$, $v_1 = g$.

This result coupled with lemma 2 enables us to verify the tentative result (1.5): that is, for the local solution $u(x,t)$ the energy functional $E(u)$ defined by (1.3) is in fact invariant throughout $[0, t_0]$, provided it exists initially. It has been established under the assumptions of lemma 2 that u satisfies the partial differential equation (1.2) pointwise in $\mathbb{R} \times [0, t_0]$, each term of the equation being bounded; hence u also satisfies the integrated form (1.4) of the equation. Upon an integration with respect to t and an application of Fubini's theorem, (1.4) gives

$$\int_{-R}^{R} (u^2+u_x^2)\, dx - \int_{-R}^{R} (g^2+g'^2)\, dx = -2\int_0^t [\tfrac{1}{2}u^2 + \tfrac{1}{3}u^3 - uu_{xt}]_{-R}^{R}\, d\tau, \tag{3.10}$$

in which the three integrals certainly exist since R is finite and by lemma 2 the integrands are all continuous functions. Now suppose that the second integral on the left-hand side remains bounded in the limit as $R \to \infty$. Since the integrand on the right-hand side is uniformly bounded on the whole of $\mathbb{R}$, (3.10) shows that the first integral on the left-hand side must also be bounded in this limit. Therefore, since the integrand is non-negative, the monotone-convergence theorem establishes that this integral exists for $R \to \infty$. But by lemma 3, if g and g' are asymptotically null, then u, u_t, u_x and u_{xt} are all asymptotically null for $0 \leqslant t \leqslant t_0$; and it follows by the dominated-convergence theorem that the right-hand side of (3.10) converges to zero as $R \to \infty$. Thus

$$\begin{aligned} E(u) = \int_{-\infty}^{\infty} (u^2+u_x^2)\, dx &= \int_{-\infty}^{\infty} (g^2+g'^2)\, dx \\ &= E_0, \quad \text{say}, \end{aligned} \tag{3.11}$$

throughout the interval $[0, t_0]$. The condition that g and g' are asymptotically null is essential here, but we note that it does not need to be assumed separately, being in fact already implied by the assumption $E_0 < \infty$ in combination with the assumption $g \in C^2(\mathbb{R})$ of lemma 2. A proof of this useful fact is presented in appendix 2.

As was noted at the end of § 1, the condition $E(u) < \infty$ implies that u is a continuous function of x whose magnitude has the uniform bound $E^{\frac{1}{2}}$ (see (1.6)). Thus it appears from (3.11) that $u(x, t_0)$ provides the same set of properties that, when assumed for $u(x, 0)$, enabled the existence of a solution to be proved for $0 \leqslant t \leqslant t_0$. Hence the argument can be repeated any number of times, extending without limit the time-interval over which a solution is guaranteed; and accordingly our main result is established as follows:

THEOREM 1. Let $g(x)$ satisfy the conditions

(i) $\int_{-\infty}^{\infty} (g^2 + g'^2)\,\mathrm{d}x = E_0 < \infty$,

(ii) $g \in C^2(\mathbb{R})$,

which imply that g and g' are asymptotically null (see appendix 2). Then the partial differential equation

$$u_t + u_x + uu_x - u_{xxt} = 0$$

has a solution $u \in \mathscr{C}^{2,\infty}_{\infty}$ which satisfies

$$u(x, 0) = g(x).$$

At any $t \in [0, \infty)$ the functions u, u_x and all their derivatives with respect to t are asymptotically null, and consequently $E(u) = E_0$.

Comments concerning function spaces

We now introduce some further definitions and notation which are needed for the generalizations to be discussed immediately below (as well as for § 4 and appendix 1), but which also furnish some commentary on the preceding result. As customary, the Banach space of (equivalence classes of) functions $w(x)$ that are square integrable on $\mathbb{R}$ is denoted by $L_2(\mathbb{R})$, and the norm of this space is written $\|w\|_2$. We have already mentioned in § 1 the so-called Sobolev space $W_2^1(\mathbb{R})$, consisting of L_2 functions with generalized first derivatives that are also L_2 functions, and we write for the norm of this space

$$\|w\|_{1,2} = (\|w\|_2^2 + \|w_x\|_2^2)^{\frac{1}{2}},$$

which is the same as $E^{\frac{1}{2}}(w)$ [cf. (1.3)]. A fact already used in this section of the paper and in § 2 is that $W_2^1(\mathbb{R}) \subset C(\mathbb{R})$.

We shall be required to consider the collection of functions $f(x, t)$ defined on $\mathbb{R} \times [0, T]$ such that $f \in L_2(\mathbb{R})$ for each $t \in [0, T]$, and such that the correspondence $t \to f(x, t)$ is a continuous mapping of $[0, T]$ into $L_2(\mathbb{R})$. This becomes a Banach space, say $\mathscr{L}_T$, under the norm

$$\|f\|_{\mathscr{L}} = \sup_{0 \leqslant t \leqslant T} \|f(x, t)\|_2. \tag{3.12}$$

Consideration also needs to be given to the collection of functions $v(x, t)$ defined on $\mathbb{R} \times [0, T]$ which are continuous in t, and which for each $t \in [0, T]$ are elements of $W_2^1(\mathbb{R})$, being therefore continuous functions of x too. This becomes a Banach space, say $\mathscr{W}_T$, under the norm

$$\|v\|_{\mathscr{W}} = \sup_{0 \leqslant t \leqslant T} \|v(x, t)\|_{1,2}. \tag{3.13}$$

When specifically allowed, the domain of definition may extend to the whole t-axis: i.e. $[0, T]$ becomes $[0, \infty)$.

It is noteworthy that a local existence theorem virtually equivalent to lemma 1 can be established almost as readily by considering the operator A to act in the space $\mathscr{W}_{t_0}$, instead of in the space $\mathscr{C}_{t_0}$ considered in the given proof of lemma 1 (see appendix 1). The assumption

(3.2) of the lemma has to be replaced by $\|g\|_{1,2} = b < \infty$ – as indeed was done in using the lemma to obtain our global result, theorem 1. Sufficient conditions for A to be a contractive mapping of a ball can be found in just the same form as (3.4) and (3.5).

By appeal to the foregoing ideas the result stated in theorem 1 may be interpreted in another way, which is perhaps more satisfying mathematically. The conditions (i) and (ii) of the theorem mean that the initial waveform $g(x)$ is drawn from the intersection of the two spaces $W_2^1(\mathbb{R})$ and $C^2(\mathbb{R})$, whereas the solution $u(x, t)$ belongs to the intersection of $\mathscr{W}_\infty$ and the space $\mathscr{C}_\infty^{2,\infty}$ noted in the statement of the theorem. Thus the process of solving the initial-value problem may be regarded as a mapping from the first into the second of these intersections.† This notion is illustrated in the following diagram:

INITIAL WAVEFORMS: $g(x) \in W_2^1(\mathbb{R}) \cap C^2(\mathbb{R})$,

$\downarrow$

SOLUTIONS: $u(x, t) \in \mathscr{W}_\infty \cap \mathscr{C}_\infty^{2,\infty}$.

It can be shown (see § 4) that the mapping represented here by the vertical arrow is unique and continuous, which, coupled with our existence theorem, establishes that the considered problem is well set in the sense defined by Hadamard (1923).

The effect of a forcing term

An important generalization of the present problem is provided by the equation

$$u_t + u_x + uu_x - u_{xxt} = f(x, t), \tag{3.14}$$

in which $f(x, t)$ is a prescribed function. In applications this function may represent some kind of forcing action on the physical system, whose response evolving from a given initial state is described by the solution u. From a theoretical standpoint, moreover, the study of equation (3.14) is important in that the right-hand side may also be considered to represent the net error entailed in equation (1.2), (2.19) as an approximate model for some particular wave system. The error will evidently take the form of some functional transformation of the dependent variable u [cf. the discussion in the paragraph following (2.15)], but an appraisal of its effect may be made by treating it implicitly as a direct function of x and t as indicated in (3.14): in obvious respects the following analysis is still applicable if f depends on x and t through dependence on $u(x, t)$. As was discussed in § 2, the error has a small parameter ϵ as a factor, but its formal smallness for $\epsilon \ll 1$ is inadequate justification for the approximate model unless (3.14) can be shown still to have strong solutions for an appropriately general right-hand side. We shall discuss the theory of equation (3.14) in outline only, since in many details it parallels the arguments leading to theorem 1.

We assume that $f(x, t)$ is defined on $\mathbb{R} \times [0, T]$, for a given finite $T > 0$, and satisfies the conditions

$$\text{(i)}\quad f \in \mathscr{C}_T, \qquad \text{(ii)}\quad f \in \mathscr{L}_T. \tag{3.15}$$

Condition (i) is evidently necessary for (3.14) to have classical solutions in the sense we have adopted, but it could be relaxed if only weak solutions were in question. Condition (ii) is, as we

† The first intersection, which is between two Banach spaces, can itself be made into a Banach space under a suitable composite norm (e.g. the sum of the norms respective to each of the intersecting spaces). Although $\mathscr{C}^{2,\infty}$ is not a Banach space, we get one if ∞ in the definition of the space is relaxed to any finite integer. In fact the space $\mathscr{C}^{2,1}$ is sufficiently restricted to delimit classical solutions of the partial differential equation (1.2): the further differentiability of solutions with respect to t, as asserted by theorem 1, is a bonus without crucial significance. Subject to such a restriction of definition the second intersection considered above can also itself be made into a Banach space.

shall see, a limitation on the 'energy' imparted by the forcing action that f represents. Note that an alternative to (i) and (ii) is the single condition $f \in \mathscr{W}_T$.

Counterparts to lemmas 1 and 2 can be established on the same lines as before. In place of the integral equation (3.1), we obtain from (2.14)

$$u = Au + Cf, \tag{3.16}$$

where

$$Cf = \int_0^t \int_{-\infty}^{\infty} \tfrac{1}{2} e^{-|x-\xi|} f(\xi, \tau) \, d\xi \, d\tau. \tag{3.17}$$

Either of the conditions (i) or (ii) of f ensures that Cf is an element of the Banach space $\mathscr{C}_{t_0}$ considered in the proof of lemma 1 (where without loss of generality we can assume that $t_0 < T$), and we may use either of the estimates

$$\|Cf\|_{\mathscr{C}} \leqslant t_0 \|f\|_{\mathscr{C}},$$

or

$$\|Cf\|_{\mathscr{C}} \leqslant \tfrac{1}{2} t_0 \|f\|_{\mathscr{L}},$$

the second of which follows from (3.17) by virtue of the Schwarz inequality. Again making the assumption (3.2) of lemma 1 and adapting the previous argument, we easily find sufficient conditions for the transformation $Cf + A$ on the right-hand side of (3.16) to be a contractive mapping of the ball $\|v\|_{\mathscr{C}} \leqslant R$: specifically, the conditions are (3.4) as before and, in place of (3.5),

$$b + \|Cf\|_{\mathscr{C}} \leqslant (1-\theta) R.$$

Hence the existence of a continuous solution of (3.15) throughout a sufficiently small interval $[0, t_0]$ follows directly. The solution is the strong limit of the Picard sequence generated by the formula

$$v_n = Av_{n-1} + Cf, \quad v_1 = g.$$

The required regularity properties of the solution may be demonstrated as in the proof of lemma 2. We find that if $g \in C^2(\mathbb{R})$ and the condition (i) in (3.15) is satisfied, then u_x, u_{xx} and their derivatives with respect to t exist as continuous functions, and hence it can be confirmed that u is a classical solution of the partial differential equation (3.14).

It remains to obtain an *a priori* estimate of $E(u)$, by means of which, as before, the local existence result can be extended over the whole of the interval $[0, T]$. We first note that, subject to the condition (ii) in (3.15), Cf is an asymptotically null function of x. A simple demonstration of this fact is provided by interpreting Cf as a convolution over $\mathbb{R}$ and thus recognizing that it is the (inverse) Fourier transform of the product

$$\frac{1}{1+k^2} \int_0^t \hat{f}(k, \tau) \, d\tau.$$

Since $f \in \mathscr{L}_T$, the Fourier transform $\hat{f}(k, t)$ is an L_2 function of k depending continuously on t, and hence so is the integral in the preceding expression. Being thus the product of two L_2 functions, this expression is an L_1 function and therefore its transform Cf vanishes as $x \to \pm\infty$, according to the Riemann–Lebesgue theorem for integrals (Rudin 1966, § 9.6). It similarly appears that $(Cf)_x$, $(Cf)_t$ and $(Cf)_{xt}$ are asymptotically null. Hence, by reasoning as in the proof of lemma 3, we may conclude that, provided g and g' are asymptotically null, the local solution u together with u_x, u_t and u_{xt} are all asymptotically null. Using these facts after multiplying (1.2) by u and integrating with respect to x, as in the derivation of (1.5), we find that

$$\frac{1}{2} \frac{dE(u)}{dt} = \int_{-\infty}^{\infty} uf \, dx, \tag{3.18}$$

provided $E(u)$ exists. That this last condition is satisfied by the local solution can be verified by considering the limit of $E(v_n)$, where $\{v_n\}$ is the Picard sequence starting with $v_1 = g$ (so that $E(v_1) = E_0$) whose strong limit in $\mathscr{C}_{t_0}$ is the solution; but we may suitably pass over the details. From (3.18) there follows, by use of the Schwarz inequality,

$$\frac{1}{2}\frac{\mathrm{d}E(v)}{\mathrm{d}t} \leqslant \|u\|_2 \|f\|_{\mathscr{L}} \leqslant E^{\frac{1}{2}}(u)\,\|f\|_{\mathscr{L}},$$

from which we infer that

$$E^{\frac{1}{2}}(u) \leqslant E_0^{\frac{1}{2}} + \|f\|_{\mathscr{L}}\, t \tag{3.19}$$

for $t \geqslant 0$. Since the supremum of $|u|$ is bounded from above by $E^{\frac{1}{2}}(u)$ [cf. (1.6)], we thus have

$$\sup_{\substack{x\in\mathbb{R}\\ t\in[0,T]}} |u(x,t)| \leqslant E_0^{\frac{1}{2}} + \|f\|_{\mathscr{L}}\, T, \tag{3.20}$$

which can serve as the required estimate.

The fixed-point argument establishing local existence can now, as before, be applied repeatedly, so that we arrive at the following result.

THEOREM 2. Let $g(x)$ satisfy the conditions (i) and (ii) of theorem 1, and let $f(x,t)$ satisfy the conditions (3.15) for a given finite $T > 0$. Then the partial differential equation (3.14) has a solution $u(x,t) \in \mathscr{C}_T^{2,1} \cap \mathscr{W}_T$ which satisfies $u(x,0) = g(x)$.

Another type of initial condition, arising in problems of bore propagation

There is an important range of applications for long-wave models that is not covered by theorem 1. In these problems the initial waveform converges to zero for $x \to \infty$, but converges to a non-zero constant for $x \to -\infty$. In particular, when this constant is positive, the model simulates the evolution of a bore, or positive surge, such as propagates along a uniform open channel subsequent to the withdrawal of a partition separating stretches of water with different surface levels. The numerical solutions of the regularized equation computed by Peregrine (1964), as mentioned in § 2, were in fact for just such an application. Theorem 1 is unavailing in this case, of course, because the initial waveform is not an L_2 function (i.e. condition (i) of the theorem is not satisfied). However, a slight modification of the argument establishing theorem 1 leads to a comparable result for this case.

Let us assume that the initial waveform, say $h(x)$, satisfies the conditions†

$$\left.\begin{array}{ll}\text{(i)}\quad h\in C^2(\mathbb{R}), & \text{(ii)}\quad h'\in L_2(\mathbb{R}),\\ \text{(iii)}\quad h, h' \to 0 \quad \text{as} \quad x\to\infty, & \\ \text{(iv)}\quad h \to \text{const.}\ (\neq 0),\quad h'\to 0 \quad \text{as} \quad x\to -\infty. & \end{array}\right\} \tag{3.21}$$

We first note that by virtue of condition (i) our previous conclusions concerning local existence (lemma 1) and regularity of solutions (lemma 2) still hold. But a new line of argument is needed to establish a global upper bound on $|u|$, where u is the solution of (1.2) satisfying $u(x,0) = h(x)$.

Substituting $u(x,t) = h(x) + v(x,t)$, we obtain from (1.2)

$$(v - v_{xx})_t + (h + v + \tfrac{1}{2}h^2 + hv + \tfrac{1}{2}v^2)_x = 0.$$

After multiplication by v, this is next integrated with respect to x between $-\infty$ and ∞. On the

† The condition, included in (iii) and (iv), that h' is asymptotically null is actually superfluous. According to the lemma proved in appendix 2, this condition is implied by (i) and (ii).

assumptions that $E(v)$ exists and that v, v_x, v_t, v_{xt} are asymptotically null, it follows after two terms have been integrated by parts that

$$\frac{1}{2}\frac{\mathrm{d}E(v)}{\mathrm{d}t}+\int_{-\infty}^{\infty}\{(1+h)\,h_x v - hvv_x\}\,\mathrm{d}x = 0 \tag{3.22}$$

[cf. (1.4) and (1.5)]. The assumptions leading to (3.22) can now be verified in much the same way as with regard to (3.18), and again we may suitably pass over the details.

Writing
$$H = \sup_{x\in\mathbb{R}} |h(x)|,$$
we infer from (3.22) that

$$\frac{1}{2}\frac{\mathrm{d}E(v)}{\mathrm{d}t} \leqslant (1+H)\int_{-\infty}^{\infty} |h_x|\,.\,|v|\,\mathrm{d}x + H\int_{-\infty}^{\infty} |v|\,.\,|v_x|\,\mathrm{d}x$$
$$\leqslant (1+H)\,\|h_x\|_2\,\|v\|_2 + H\,\|v\|_2\,\|v_x\|_2$$

by the Schwarz inequality. Here $\|v\|_2$ and $\|v_x\|_2$ are functions of time, and obviously $\|v\|_2 \leqslant E^{\frac{1}{2}}(v)$ and $\|v\|_2\,.\,\|v_x\|_2 \leqslant \frac{1}{2}E(v)$. Therefore

$$\frac{1}{2}\frac{\mathrm{d}E(v)}{\mathrm{d}t} \leqslant l(1+H)\,E^{\frac{1}{2}}(v) + \tfrac{1}{2}HE(v),$$

where we have written $l = \|h_x\|_2$. From this differential inequality and the fact that $v(x, 0) \equiv 0$, it follows that

$$E^{\frac{1}{2}}(v) \leqslant 2l(1+H^{-1})\,(\mathrm{e}^{\frac{1}{2}Ht}-1) \tag{3.23}$$

for $t \geqslant 0$. The supremum of $|v|$ is bounded from above by $E^{\frac{1}{2}}(v)$, and so, given any finite $T > 0$, we have that

$$\sup_{\substack{x\in\mathbb{R}\\ t\in[0,\,T]}} |v(x,t)| \leqslant 2l(1+H^{-1})\,(\mathrm{e}^{\frac{1}{2}HT}-1). \tag{3.24}$$

An upper bound on the magnitude of the solution $u(x, t)$ now follows from

$$\sup|u| = \sup|h+v| \leqslant H + \sup|v|.$$

Hence, essentially as before, the argument establishing local existence can be applied repeatedly, leading to the following global result.

THEOREM 3. Let $h(x)$ satisfy the conditions (3.21). Then, for any finite $T > 0$, equation (1.2) has a solution $u(x, t) \in \mathscr{C}_T^{2,\infty}$ which satisfies $u(x, 0) = h(x)$, and which is such that $(u-h) \in \mathscr{W}_T$.

Analyticity of solutions with respect to t

It was shown in the proof of lemma 2 that a solution $u(x, t)$ of (3.1) defined on $\mathbb{R}\times[0,\infty)$ is a C^∞ function of t. In keeping with this conclusion, we shall now prove that the solution guaranteed by theorem 1 is in fact analytic in t. That is, for any finite x_1 and $t_1 \in (0, \infty)$, we consider the formal Taylor series

$$U(t) = \sum_{m=0}^{\infty} [\partial_t^m u(x_1, t)]_{t=t_1} \frac{(t-t_1)^m}{m!}, \tag{3.25}$$

and show that $U(t)$ converges to $u(x_1, t)$ in some neighbourhood of t_1. Proof is achieved by confirming that the Lagrange remainder in Taylor's theorem tends to zero as $m \to \infty$.

The successive derivatives of u with respect to t are given by the formula (3.7), and we note that $\partial_t^{m-1}(u+\frac{1}{2}u^2)$ appearing on the right-hand side is just a polynomial in $u, u_t, \ldots, \partial_t^{m-1}u$, all of whose coefficients are positive. We also observe with regard to (3.7) that

$$\int_{-\infty}^{\infty} |K(x_1-\xi)|\,\mathrm{d}\xi = 1.$$

Hence, if $C_m(u)$ is taken to denote a uniform (over $\mathbb{R}$) upper bound on the magnitude of the mth derivative of u with respect to t, the formula (3.7) is seen to provide a succession of such bounds related by

$$C_m(u) = C_{m-1}(u+\tfrac{1}{2}u^2) \quad (m = 1, 2, \ldots).$$

Thus, given an upper bound $C_0 > 0$ on $|u|$ at $t = t_1$, we can compute a simultaneous upper bound on $|\partial_t^m u|$ for any m; and the procedure for this is identical with the calculation of the mth derivative at $t = t_1$ of the function $V(t)$ determined by

$$\mathrm{d}V/\mathrm{d}t = V + \tfrac{1}{2}V^2, \quad V(t_1) = C_0 > 0. \tag{3.26}$$

If moreover the bound C_0 applies uniformly in t, then, since the derivatives of V are evidently all positive at $t = t_1$ if $V(t_1) > 0$, we have for all $m > 0$

$$\sup |\partial_t^m u(x_1, t)| \leqslant V^{(m)}(t_1),$$

where the supremum is taken over a neighbourhood of t_1. Hence convergence of the Taylor series for $V(t)$ about $t = t_1$ implies evanescence of the Lagrange remainder in respect of the series $U(t)$.

From the last statement it follows that, if C_0 is given and the $V(t)$ accordingly defined by (3.26) is analytic, then *a fortiori* $U(t)$ is analytic. But under the conditions of theorem 1 an explicit upper bound on $|u|$ applying uniformly in $\mathbb{R}\times[0,\infty)$ is available, namely $C_0 = E^{\frac{1}{2}}(u) = E_0^{\frac{1}{2}}$. And the analyticity of $V(t)$ in a neighbourhood of t_1 can be confirmed immediately. We find directly from (3.26) that, provided $t - t_1 < \ln\{(2+C_0)/C_0\}$,

$$V(t) = \sum_{m=0}^{\infty} V^{(m)}(t_1)\,\frac{(t-t_1)^m}{m!} = \frac{2C_0}{(2+C_0)\,\mathrm{e}^{t_1-t} - C_0}.$$

This completes the demonstration that the solution assured by theorem 1 is analytic in t.

4. Uniqueness and stability

In this section two other favourable aspects of the regularized long-wave equation are demonstrated. We first show that the solution corresponding to any given initial waveform is unique. We then consider two 'stability' properties concerning the relationship between solutions that correspond to different initial values and to different forcing functions $f(x,t)$ as in (3.14). It appears that, over any finite time-interval, solutions depend continuously on each of these determining factors; which conclusion, together with the uniqueness result, complements the existence theory of § 3 in establishing that the considered problem is well set in the sense first explained by Hadamard (1923). Here again we recall Hadamard's conception that three requirements – existence, uniqueness and the continuous dependence of solutions on prescribed influences – are generally essential to the practical utility of mathematical models for evolutionary processes. The aforesaid restriction to a finite time-interval appears quite natural for the present problem. In respect of varied initial values the continuity of solutions over an *unbounded* time-interval is evidently out of the question, as witness a comparison between solutions representing two nearly equal solitary waves each of which is initially centred on the same point in $\mathbb{R}$. However slight the difference in the amplitudes of the waves, the corresponding difference in their velocities of propagation destines that at sufficiently large times these solutions will be far apart, judged by any of the measures of closeness adopted here.

As was noted in the context of (3.14), the continuous dependence of solutions on the forcing function is particularly significant with regard to the validation of the regularized equation (1.2) as an approximate model for long waves. We may consider the exact behaviour of the physical system in question to be represented implicitly by an equation in the form (3.14), where now the term on the right-hand side stands for the remainder that is approximated by zero when we take (1.2) as a model equation. For small remainders, the approximation is justified by the property that solutions depend continuously on them, which is implied if solutions are shown to depend so in general on the forcing term.

Uniqueness

Let u_1 and u_2 be two solutions of equation (1.2) which have the properties assured by theorem 1; or let them be two solutions of equation (3.14), with a given forcing term f, which have the properties assured by theorem 2. In either case, subtracting the equation for u_2 from that for u_1 and writing $w = u_1 - u_2$, we obtain

$$w_t + w_x + \tfrac{1}{2}\{(u_1+u_2)\, w\}_x - w_{xxt} = 0. \tag{4.1}$$

The properties of u_1 and u_2 imply that each term in (4.1) is continuous in x and bounded on $\mathbb{R}$. Hence, when we multiply (4.1) by w and integrate with respect to x over the interval $[-R, R]$, integrations by parts of the third and fourth terms are permissible, leading to

$$\int_{-R}^{R} (ww_t + w_x w_{xt})\, \mathrm{d}x - \frac{1}{2}\int_{-R}^{R} (u_1+u_2)\, ww_x \mathrm{d}x + [\tfrac{1}{2}w^2 + \tfrac{1}{2}(u_1+u_2)\, w^2 - ww_{xt}]_{-R}^{R} = 0. \tag{4.2}$$

Since u_1, u_2, $(u_1)_{xt}$ and $(u_2)_{xt}$ are asymptotically null, so are w and w_{xt}. Thus the integrated terms in (4.2) vanish in the limit as $R \to \infty$. The second integral evidently converges in this limit since, for each finite $t \geqslant 0$, u_1 and u_2 are bounded on $\mathbb{R}$ and w, like u_1 and u_2, is an element of $W_2^1(\mathbb{R})$. In fact we have

$$\left|\int_{-\infty}^{\infty} (u_1+u_2)\, ww_x \mathrm{d}x\right| \leqslant \tfrac{1}{2}C(t)\, \|w\|_{1,2}^2,$$

where

$$C(t) = \sup_{x\in\mathbb{R}} |u_1+u_2| \leqslant \|u_1\|_{1,2} + \|u_2\|_{1,2}. \tag{4.3}$$

It follows that the first integral in (4.2) also converges in the limit as $R \to \infty$; and, from the fact that $\|u_1\|_{1,2}$, $\|u_2\|_{1,2}$ and hence $\|w\|_{1,2}$ are bounded over any finite time-interval, we may infer that

$$\int_{-\infty}^{\infty} (ww_t + w_x w_{xt})\, \mathrm{d}x = \frac{1}{2}\frac{\mathrm{d}E(w)}{\mathrm{d}t},$$

where $E(w) = \|w\|_{1,2}^2$. Thus (4.1) leads to the differential inequality

$$\frac{\mathrm{d}E(w)}{\mathrm{d}t} \leqslant \tfrac{1}{2}C(t)\, E(w), \tag{4.4}$$

which implies that

$$E(w) \leqslant [E(w)]_{t=0} \exp\left(\frac{1}{2}\int_0^t C(\tau)\, \mathrm{d}\tau\right). \tag{4.5}$$

Now, if u_1 and u_2 correspond to the same initial waveform $g(x)$, so that w is identically zero at $t = 0$, then $E(w) = 0$ at $t = 0$. Hence (4.5) shows that $E(w) = 0$, and therefore $u_1 \equiv u_2$, for all finite $t > 0$.

This conclusion can be rephrased as follows:

THEOREM 4. The solution of equation (1.2) guaranteed by theorem 1 is *unique*, as also is the solution of the generalized equation (3.14) guaranteed by theorem 2.

By a straightforward adaption of the preceding argument it can be shown that the solution of (1.2) guaranteed by theorem 3, for an initial waveform that does not vanish as $x \to -\infty$, is also unique.

Stability

The result (4.4) may be used further to demonstrate the continuity of solutions with respect to varied initial values. Let u_1 and u_2 be the unique solutions of (3.14) [or, in particular, of (1.2)] such that $u_1(x, 0) = g_1(x)$ and $u_2(x, 0) = g_2(x)$. Write $\Delta g = g_1 - g_2$. Considering the solutions as elements of the Banach space $\mathscr{W}_T$ defined in § 3, we propose that $\|w\|_{\mathscr{W}} = \|u_1 - u_2\|_{\mathscr{W}}$ can be made arbitrarily small by taking g_2 close enough to g_1, in the sense that $\|\Delta g\|_{1,2}$ is made small enough.

Suppose that
$$\|\Delta g\|_{1,2} \leqslant \delta.$$
Then, by means of the inequality (3.20) used in the proof of theorem 2, an upper bound for $C(t)$, as defined by (4.3), may be obtained in the form
$$\begin{aligned}\max_{0\leqslant t\leqslant T} C(t) &\leqslant \|g_1\|_{1,2} + \|g_2\|_{1,2} + 2\|f\|_{\mathscr{L}} T \\ &\leqslant 2\|g_1\|_{1,2} + 2\|f\|_{\mathscr{L}} T + \delta, \\ &= c, \quad \text{say,}\end{aligned} \tag{4.6}$$
which is independent of g_2. The result (4.5) now tells us that
$$\begin{aligned}\|w\|_{\mathscr{W}} = \sup_{0\leqslant t\leqslant T} \|w\|_{1,2} &\leqslant \|\Delta g\|_{1,2} \exp\left(\tfrac{1}{4}cT\right) \\ &\leqslant \delta \exp\left(\tfrac{1}{4}cT\right),\end{aligned} \tag{4.7}$$
and this obviously establishes the proposition in question. (This conclusion will be interpreted in another way presently, as part of a wider proposition.)

We next consider continuity with respect to varied forcing functions. Suppose that the forcing function is changed from f_1 to f_2, and respectively the unique solutions of (3.14) are u_1 and u_2. Write $\Delta f = f_1 - f_2$. As in theorem 2, it is assumed that the functions f_1 and f_2 are elements of the Banach space $\mathscr{L}_T$, which means that for each $t \in [0, T]$ they are elements of $L_2(\mathbb{R})$ (and moreover that the correspondences $t \to f_1(x, t)$ and $t \to f_2(x, t)$ are continuous mappings of $[0, T]$ into $L_2(\mathbb{R})$). Proceeding as before by subtracting the equation for u_2 from that for u_1, multiplying by $w = u_1 - u_2$ and integrating with respect to x, we obtain in place of (4.4)
$$\mathrm{d}E(w)/\mathrm{d}t \leqslant \tfrac{1}{2}C(t)\,E(w) + 2\left|\int_{-\infty}^{\infty} w\,\Delta f\,\mathrm{d}x\right|$$
for $0 \leqslant t \leqslant T$. By use of the Schwarz inequality and the definition (3.12) of the norm in $\mathscr{L}$, this leads to
$$\mathrm{d}E(w)/\mathrm{d}t \leqslant \tfrac{1}{2}C(t)\,E(w) + 2\|\Delta f\|_{\mathscr{L}} E^{\frac{1}{2}}(w). \tag{4.8}$$
Supposing that
$$\|\Delta f\|_{\mathscr{L}} \leqslant \eta,$$
and allowing as before that u_1 and u_2 may have different initial values, we can write down an upper bound for $C(t)$ akin to (4.6), namely
$$C(t) \leqslant c = 2\|g_1\|_{1,2} + 2\|f_1\|_{1,2} T + \delta + \eta. \tag{4.9}$$
Hence it follows straightforwardly from (4.8) that
$$\|w\|_{\mathscr{W}} \leqslant \delta \exp\left(\tfrac{1}{4}cT\right) + 4\eta c^{-1}\{\exp\left(\tfrac{1}{4}cT\right) - 1\}, \tag{4.10}$$
and the right-hand side of this inequality is independent of f_2 and g_2. Thus, if $\delta = 0$ (i.e. u_1 and u_2 have the same initial values), $u_1 - u_2$ can be made arbitrarily small by keeping f_2 sufficiently close to f_1, so that η is small enough. This property is precisely what we mean by the continuity of solutions with respect to varied forcing functions.

The case of *simultaneously* varied initial values and forcing functions is also covered by the result (4.10). Accordingly, to complete the discussion, we formulate a more refined statement of stability properties which generalizes the foregoing conclusions about the continuity of solutions. We consider the topological product
$$\mathscr{G} = W_2^1(\mathbb{R}) \cap C^2(\mathbb{R}) \times \mathscr{L}_T \cap \mathscr{C}_T,$$

any element of which has the form (g,f), where $g(x) \in W_2^1(\mathbb{R}) \cap C^2(\mathbb{R})$ and $f(x,t) \in \mathscr{L}_T \cap \mathscr{C}_T$. In the usual way for such products, $\mathscr{G}$ is given a Banach-space structure by defining its norm to be

$$\|(g,f)\|_{\mathscr{G}} = \|g\|_{W_2^1 \cap C^2} + \|f\|_{\mathscr{L} \cap \mathscr{C}}. \tag{4.11}$$

[For explanation of the terms on the right-hand side of (4.11), see the subsection *Comments concerning function spaces* in § 3.] Since g and f as delimited here satisfy the hypotheses of theorem 2, we can associate with any element $(g,f) \in \mathscr{G}$ the unique solution of (3.14) that has the particular g as initial waveform and the particular f as forcing function.

The general proposition in view may now be stated as follows:

THEOREM 5. The mapping $\mathbf{U}: \mathscr{G} \to \mathscr{W}$, which assigns to each element $(g,f) \in \mathscr{G}$ the corresponding unique solution u of the regularized equation (3.14), is continuous.

The proof is obvious in the light of (4.9) and (4.10), being a simple extension of the argument used previously. Note that if

$$\|(g_1,f_1) - (g_2,f_2)\|_{\mathscr{G}} \leqslant \iota, \tag{4.12}$$

then $\delta \leqslant \iota$ and $\eta \leqslant \iota$, where δ and η are the bounds on $\|\Delta g\|_{1,2}$ and $\|\Delta f\|_{\mathscr{L}}$ considered previously. Hence it appears that, given any particular element (g_1,f_1) and any (small) $\epsilon > 0$, a value $\iota > 0$ can be chosen to ensure that

$$\|u_1 - u_2\|_{\mathscr{W}} \leqslant \epsilon$$

for all (g_2,f_2) satisfying (4.12). This property amounts to continuity of the mapping $\mathbf{U}$ as stated in the theorem.

Since theorem 2 asserts that the guaranteed solutions of (3.14) lie in the intersection $\mathscr{W}_T \cap \mathscr{C}_T^{2,1}$, it is relevant to ask whether the correspondence between $\mathscr{G}$ and the class of solutions is also continuous considered as a mapping of $\mathscr{G}$ into this intersection. In fact, an affirmative answer to this question can be found by further study of equation (4.1), but we pass over this aspect here.

5. CONCLUSION

In the foregoing exposition two distinct lines of inquiry were pursued with a common aim, namely the advocacy of equation (1.2) as a preferable long-wave model in applications where the Korteweg–de Vries equation (1.1) is an alternative. First, in § 2, relying on descriptive rather than rigorous mathematics, we examined the general origins of these equations in physical problems, and we identified the elements that appear to correlate the numerous applications already found for the KdV equation. The leading point of the discussion in § 2 was that our regularized equation is formally equivalent to the KdV equation, considered in its original perspective as an approximation accounting only to first order for small nonlinear and dispersive effects. In the last part of § 2 various problematical aspects of the KdV equation were pointed out, all of which arise from spuriously potent short-wave behaviour determined by the equation, and none of which is posed by the regularized equation whose short-wave properties are comparatively feeble as befits a long-wave model.

Secondly, in §§ 3 and 4, a basic theory for equation (1.2), and its extension (3.14) with a forcing term, was worked out, showing that the equation has wholly satisfactory mathematical properties. In particular, as established in § 4, the stability of solutions with respect to variations in prescribed influences (initial values and forcing functions) is a decidedly favourable property as regards applications of the model. The initial-value problem for (1.2) has in principle been solved in §§ 3 and 4, subject to assumptions much weaker than anything so far claimed for the KdV equation. Offsetting these advantages, however, it seems that for (1.2) there is no counterpart to many of the remarkable formal results obtained for the KdV equation in recent years. For example, whereas infinitely many invariant functionals in the form (2.21) exist for solutions of the KdV equation (Miura *et al.* 1968), a determined search by the present authors revealed only three for solutions of (1.2), including the linear invariant obviously obtainable by integrating (1.2) from $x = -\infty$ to $x = \infty$ [see remarks at the end of § 1]. But this appears in no respect a serious shortcoming in the theory of equation (1.2). The single invariant property (1.5) suffices for the purposes of the existence theory; and, as we suggested in § 2, the KdV results may be somewhat illusive, depending as they do on assumptions about the regularity and asymptotic properties of solutions that cannot be justified in any straightforward way. Also, the provision of just three invariants by the model seems perfectly reasonable in some applications, such as to water waves, for they then correspond to conservation of mass, energy and momentum in the original physical system.

It should be recognized that the validity of equation (1.2) as a long-wave approximation applying to particular physical systems has not been proved here. Indeed this issue is beyond the scope of any general discussion like that in § 2 because, obviously, it will turn on the complete specifications of the system in question. That is, the exact dynamical equations for the system will need to be considered in order to formulate estimates that might verify (1.2) as an approximation. In demonstrating the mathematical expedience of (1.2) and (3.14), however, we have established a basis which would provide readily for the complete validation of this model in particular applications. In particular, as was discussed in the context of (3.14) and in the second paragraph of § 4, our result that solutions depend continuously on the forcing function f, when it is varied within a general class of functions, would be the crucial implement in a proof of validity.

Among prospective extensions of the present work, the most immediately promising is the completion of an existence and stability theory for the generalized version of equation (1.2) that is explained in appendix 1. In this equation general dispersive properties are represented by an abstract pseudo-differential operator. For a restricted class of such operators, an existence theory may be developed by a straightforward adaption of the arguments used in § 3, and this is sketched in appendix 1; but a satisfactorily comprehensive theory, covering examples that have already been derived from significant physical problems, calls for methods transcending those applied in §§ 3 and 4.

Finally, we mention another form of model equation which can be treated by an extension of present methods. This is comparable with the model equation (2.26) proposed by Whitham (1967), which has linear terms representing *exactly* the dispersive effects suffered by infinitesimal waves, and which includes a nonlinear term derived from the first-order approximation for long waves of small but finite amplitude. It was pointed out in the context of (2.26) that Whitham's model has inexpedient short-wave properties; but this difficulty is obviated by the equation

$$u_t + L(u_x + uu_x) = 0, \tag{5.1}$$

in which as before L is the exact dispersion operator given by linearized theory, corresponding to the exact relation $c = c(k)$ between the phase velocity and wavenumber of sine waves. Since $c(k)$ is a continuous function and $c(0) = 1$, L reduces approximately to the identity operator for long waves, so that as a long-wave approximation (5.1) has the same formal standing as the other model equations that we have considered. A property generally provided by the physical problem is that $c(k)$ vanishes for $|k| \to \infty$, which implies that L is a smoother operator than identity. For example, if $c(k) < \beta(1+k^2)^{-1}$, where β is a finite constant, L is found to be a continuous operator acting from $L_2(\mathbb{R})$ into $W_2^1(\mathbb{R})$. Accordingly, after an integration with respect to t, we may obtain from (5.1) an equation of essentially the same kind as (3.1), and hence the theory of (5.1) may be developed on the present lines. Corresponding to (1.5), an invariant property of solutions of (5.1) is that

$$\int_{-\infty}^{\infty} u(L^{-1}u)\,\mathrm{d}x = \text{const.}, \tag{5.2}$$

which may be used as the basis for a global extension of a local existence theorem.

We are indebted to the Science Research Council for their support of the Fluid Mechanics Research Institute. J. J. M. also wishes to express his gratitude to the University of Essex for hospitality during the period of this collaboration.

Appendix 1. Model equations representing other forms of dispersion

While being concerned with some general issues involved in the mathematical modelling of long-wave phenomena, our discussion has concentrated on the archetypal problem presenting the KdV equation (1.1) or its alternative (1.2). The applicability of these equations to long waves in nonlinear dispersive systems is by no means universal, however, and to supplement our main subject-matter a generalization will now be reviewed, extending to systems whose dispersive properties cannot be approximated in the way that was explained in § 2. The object here is just to specify the generalized problem, not to solve it, and the details of an existence and stability theory comparable with the material of §§ 3 and 4 are suitably left for a separate account.

To define the present aspect, we need first to recall some essentials of the discussion in § 2, in particular concerning the linearized dispersion equation (2.6) and approximations to it applicable to long waves. For simplicity the relevant properties are now expressed in terms of dimensionless variables free from scaling factors, and we refer to § 2 for the principles whereby these variables are defined. Thus equation (2.6), which governs the propagation of infinitesimal waves in the $+x$ direction, is reconsidered in the form

$$u_t + (Lu)_x = 0, \tag{A 1}$$

with the operator L defined by

$$\widehat{Lu} = c(k)\,\hat{u}(k,t), \tag{A 2}$$

where $c(0) = 1$ and the circumflexes denote Fourier transforms, i.e.

$$\hat{u}(k,t) = \int_{-\infty}^{\infty} \mathrm{e}^{-\mathrm{i}kx} u(x,t)\,\mathrm{d}x.$$

The function $c(k)$ in (A 2) may be interpreted as an expression for the phase velocity of infinitesimal sine waves with wavenumber k; and it is an essential attribute of the class of physical systems in question that $c(k)$ is a continuous, even and non-negative function taking its maximum value at $k = 0$. The condition $c(0) = 1$, determined by the choice of dimensionless independent variables, means that L reduces to identity in the case of extremely long waves.

In the formal derivation of the KdV equation, or of our regularized equation (1.2) or (2.19), it is assumed that the approximation

$$c(k) = 1 - k^2 \tag{A 3}$$

is valid for small values of $|k|$: that is, the curve $c = c(k)$ can be approximated in the neighbourhood of $k = 0$ by an osculating parabola. The definition (A 2) of the operator L accordingly yields the approximation $L = I + \partial_x^2$; and, as was explained in the context of (2.18) and (2.19), a long-wave approximation with the same formal standing as this is $L = I - \partial_x \partial_t$. But suppose the function $c(k)$ does not admit the small-k approximation (A 3), as when, for example, $c(k)$ has a first but not a second derivative at $k = 0$. A long-wave approximation to L may still be definable, and may be justifiable in virtually the same way as the foregoing approximation was in § 2. However, one is then presented with a *pseudo-differential* operator which, in contrast with the differential operator presented previously, has no *local* representation alternative to its basic definition in terms of Fourier transforms. In general, if for small k one has

$$c(k) - \{1 - \alpha(k)\} = o\{\alpha(k)\} \quad (\alpha(0) = 0),$$

then the long-wave approximation to L is

$$L = I - H, \tag{A 4}$$

where the operator H defined by

$$\widehat{Hv} = \alpha(k)\,\hat{v}(k) \tag{A 5}$$

is pseudo-differential if $\alpha(k)$ is not a polynomial and, as will be assumed, $\alpha(k) \to \infty$ as $|k| \to \infty$. To be fully meaningful the definition (A 5) needs further qualification, of course, because the existence of the Fourier transforms and their inverses must be assured. The function $\alpha(k)$ will be called the *symbol* of H, which is the usual term in the recent literature on pseudo-differential operators (see, for instance, Nirenberg 1971). We note the representation of H as a convolution

$$Hv = \int_{-\infty}^{\infty} h(x - \xi)\, v(\xi)\, d\xi,$$

where $h(x)$ is the generalized function (distribution) whose Fourier transform is $\alpha(k)$. From this, or from Parseval's theorem applied to the definition (A 5), it is evident that the operator H is symmetric (or, loosely, self-adjoint), thus

$$\int_{-\infty}^{\infty} uHv\, dx = \int_{-\infty}^{\infty} vHu\, dx.$$

Now if, just as in the derivation of (2.19), the zero-order equivalence of ∂_x and $-\partial_t$ is exploited, and if nonlinear effects are represented as before, we may proceed from (A 4) to infer the general model equation

$$u_t + u_x + uu_x + (Hu)_t = 0, \tag{A 6}$$

which recovers (2.19) in the special case $\alpha(k) = k^2$. Alternatively, if we do not replace $(Hu)_x$ by $-(Hu)_t$, a corresponding generalization of the KdV equation is obtained; but this presents the same sort of difficulties as the KdV equation, such as were discussed in the last part of § 2. In particular, the artificial dispersion relation corresponding to the linearized form of the generalized KdV equation is

$$c(k) = 1 - \alpha(k),$$

giving unbounded negative phase and group velocities as $|k| \to \infty$. On the other hand, the artificial dispersion relation corresponding to (A 6) is

$$c(k) = \frac{1}{1 + \alpha(k)},$$

which, if $\alpha(k)$ is a non-decreasing function of k, has satisfactory properties like those pointed out with regard to (2.25).

Examples

To illustrate the definition of the pseudo-differential operator H, we give two examples deriving from definite physical problems. The first was shown by Benjamin (1967*a*) to arise in the theory of internal waves in a stably stratified liquid, when the prescribed variations of density are confined to a stratum whose thickness is much less than the total depth of the liquid. The operator then presented ($H = \mathscr{F}$ in the notation of the cited paper) has the symbol $\alpha(k) = |k|$, and so is comparable in 'strength' with the differential operator ∂_x – which is not symmetric. More precisely, this particular H has the property $H(Hv) = -\partial_x^2 v$. Thus it may be interpreted as the square root, in the class of real, symmetric operators, of the symmetric differential operator $-\partial_x^2$ which takes the corresponding place in the archetypal problem considered earlier.

The second example arises in the theory of axisymmetric waves in a rotating fluid, when the fluid is taken to be unbounded radially. It is assumed that the azimuthal circulation increases with radius, as is required for stability, and becomes constant either asymptotically or at the surface of a finite core. Aspects of this problem have been treated by Benjamin (1967*b*), Pritchard (1970) and Leibovich (1970), each of whom discussed a peculiar dispersive property that is consequent upon the infinite extent of the fluid. In the corresponding problem for a rotating fluid contained in a tube, the phase velocity $c(k)$ for each particular wave mode is found to be a C^∞ function of k, and the theory of moderately long waves of finite but small amplitude exemplifies the standard essentials that were explained in § 2. But when the fluid is unbounded $c(k)$ is only a C^1 function, its small-k behaviour being $c(k) - 1 \sim k^2 \ln |k|$. A model equation including this special feature was proposed by Leibovich (1970) in the generalized K dV form, with the symbol of H taken to be

$$\alpha(k) = k^2 K_0(B|k|) \quad (B > 0),$$

where K_0 is the modified Bessel function of order zero. This model correctly simulates the long-wave properties of the physical system, but it has inexpedient short-wave behaviour in that $\alpha(k)$ vanishes as $|k| \to \infty$. Thus dispersive effects disappear for very short waves, whereas nonlinear effects do not, so that the model is likely to present difficulties (particularly computational ones) of the kind that were pointed out with regard to equation (2.25). A slight change in approach, however, leads to a more expedient model in which the symbol of H has the form

$$\alpha(k) = k^2\{A + K_0(B|k|)\} \quad (A, B > 0),$$

and accordingly an explicit representation of H is

$$Hv = -\partial_x^2\left[Av(x) + \frac{1}{2}\int_{-\infty}^{\infty} \frac{v(\xi)\,d\xi}{\{(x-\xi)^2 + B^2\}^{\frac{1}{2}}}\right].$$

The strength of H, which is determined by the asymptotic behaviour of $\alpha(k)$ for large k, is in the present instance evidently the same as that of the differential operator $-\partial_x^2$.

Outline of existence theory

When allowance is made for an amply representative class of operators H, wide enough to include the two examples noted above, the theory of equation (A 6) presents intricacies considerably beyond the range of the ideas in §§ 3 and 4, and so we shall not attempt to cover them in this supplement. It is intended that equation (A 6) will be the subject of another paper. For a restricted class of operators, however, the gist of an existence theory may readily be indicated as follows.

We assume additionally that the continuous even function $\alpha(k)$ is positive for $|k| > 0$ and satisfies the asymptotic condition

$$\liminf_{|k|\to\infty} \frac{\alpha(k)}{k^2} > 0, \tag{A 7}$$

which can also be expressed, in a way bearing more directly on the subsequent argument, as

$$\frac{1}{1+\alpha(k)} = O\left(\frac{1}{k^2}\right) \quad \text{for} \quad |k| \to \infty. \tag{A 7'}$$

This condition means that $\boldsymbol{H}$ defined by (A 5) is at least as strong as the particular operator $-\partial_x^2$ that took the place of $\boldsymbol{H}$ in §§ 3 and 4. It follows from (A 7′) and from the continuity of $\alpha(k)$ that a positive constant a exists such that

$$\frac{a}{1+\alpha(k)} \leqslant \frac{1}{1+k^2}. \tag{A8}$$

As in § 3 the theory proceeds in two stages, first establishing a local (small-time) result and then a global extension. For the first stage (A 6) is converted by formal operations into

$$\begin{aligned} u &= g(x) - \int_0^t (\boldsymbol{I}+\boldsymbol{H})^{-1}(u_x + uu_x)\,\mathrm{d}\tau \\ &= g(x) + \int_0^t \boldsymbol{Z}(u+\tfrac{1}{2}u^2)\,\mathrm{d}\tau, \end{aligned} \tag{A 9}$$

where

$$\boldsymbol{Z} = \partial_x(\boldsymbol{I}+\boldsymbol{H})^{-1}$$

is the linear operator whose symbol, appearing in a definition corresponding to (A 5), is

$$\hat{Z}(k) = -\frac{\mathrm{i}k}{1+\alpha(k)}.$$

Equivalently, $\mathbf{Z}$ may be interpreted as a convolution with the function $Z(x)$ whose Fourier transform is $\hat{Z}(k)$. By appeal to Parseval's theorem and the inequality (A 8), it readily appears that $\mathbf{Z}$ is a continuous operator acting from $L_2(\mathbb{R})$ into $W_2^1(\mathbb{R})$. Hence, when the right-hand side of (A 9) is considered as a nonlinear transformation $\boldsymbol{A}u$ (as was the right-hand side of (3.1)), it can be shown that, provided $g \in W_2^1(\mathbb{R})$, $\boldsymbol{A}$ is a continuous operator acting in the Banach space $\mathscr{W}_{t_0}$ that was defined in § 3. Sufficient conditions for $\boldsymbol{A}$ to be a contractive mapping of a ball $\|v\|_{\mathscr{W}} \leqslant R$ into itself can be found which are satisfied if t_0 is small enough, and thus a solution of (A 9) is proved to exist over a small time-interval $[0, t_0]$. In the steps leading to this result, which are comparable with the proof of lemma 1, § 3, repeated use has to be made of the inequality

$$\left\| \int_0^t v(x,\tau)\,\mathrm{d}\tau \right\|_2 \leqslant \int_0^t \|v(x,\tau)\|_2\,\mathrm{d}\tau$$

(cf. Hardy, Littlewood & Polya 1952, p. 148, theorem 202).

Regularity properties of the local solution of (A 9) may be established by bootstrap arguments like those used in the proof of lemma 2, § 3. On the assumption that the initial waveform $g(x)$ is continously differentiable and also $Hg \in C(\mathbb{R})$, it can be shown that the solution is simultaneously a *strong* solution of (A 6) – in the sense that (A 6) is satisfied pointwise and each term of the equation is a continuous function. As in the case treated in § 3, the solution is found to be a C^∞ function of t.

To obtain an *a priori* estimate akin to (1.5), providing for a global extension of the existence result, the asymptotic nullity of the solution u and its derivative u_x needs to be demonstrated. On the necessary assumption that g and g' are asymptotically null, these properties can be verified from (A 9). The solution is known to be an element of $\mathscr{W}_{t_0}$, so that for each $t \in [0, t_0]$ it is a bounded and square-integrable function of x on $\mathbb{R}$. Hence $u+\frac{1}{2}u^2$ is also an L_2 function, and therefore its Fourier transform $(u+\frac{1}{2}u^2)^\wedge$ exists, being an L_2 function of the transform variable k. Now $\boldsymbol{Z}(u+\frac{1}{2}u^2)$ is the inverse Fourier transform of $\hat{Z}(k).(u+\frac{1}{2}u^2)^\wedge$, and the condition (A 7) implies that $\hat{Z}(k) \in L_2$. Being thus the product of two L_2 functions, $\hat{Z}(k).(u+\frac{1}{2}u^2)^\wedge$ is an L_1 function; and therefore the inverse transform $\boldsymbol{Z}(u+\frac{1}{2}u^2)$ is a continuous function of x which, according to the Riemann–Lebesgue theorem, vanishes at infinity. By virture of the dominated-convergence theorem, the latter property is preserved by the integral of $\boldsymbol{Z}(u+\frac{1}{2}u^2)$ with respect to t; and thus it appears from (A 9) that u is asymptotically null. Similarly, from the known fact that u_x is a L_2 function of x, so that $(1+u)\,u_x$ is also an L_2 function, it can be shown by means of (A 9) that u_x is asymptotically null.

Using the preceding conclusions after multiplying (A 6) by u and integrating with respect to x, we deduce that

$$\int_{-\infty}^{\infty} \{uu_t + u(\boldsymbol{H}u)_t\}\,\mathrm{d}x = 0.$$

In view of the fact that $\boldsymbol{H}$ commutes with ∂_t and is symmetric, it follows that

$$\int_{-\infty}^{\infty} (u^2 + u\boldsymbol{H}u)\,\mathrm{d}x = \text{const.}, \tag{A 10}$$

provided this integral exists – which property can be verified for the solution u of (A 9) if it holds for the initial waveform g. But appealing again to Parseval's theorem and the inequality (A 8), we have

$$\begin{aligned}\int_{-\infty}^{\infty} (u^2 + u\boldsymbol{H}u)\,\mathrm{d}x &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \{1+\alpha(k)\}\,|\hat{u}|^2\,\mathrm{d}k \\ &\geqslant \frac{a}{2\pi}\int_{-\infty}^{\infty} (1+k^2)\,|\hat{u}|^2\,\mathrm{d}k = a\,\|u\|_{1,2}^2.\end{aligned} \tag{A 11}$$

The combination of (A 10) and (A 11) shows that $\|u\|_{1,2}$ has an upper bound which is independent of t, being fixed by the initial data. By reasoning similar to the proof of theorem 1 in § 3, this fact may be used to complete a demonstration that equation (A 6) has a solution over an arbitrarily large time-interval.

Appendix 2. Nullity at infinity of a function class

Here we present a proof of the fact, used in § 3, that an element of $W_2^1(\mathbb{R}) \cap C^2(\mathbb{R})$ and its first derivative both converge to zero at $\pm\infty$: i.e. they are asymptotically null according to our definition. The proof will follow immediately from the following lemma which may have some independent interest.

Lemma. Let $g \in L_2(\mathbb{R}) \cap C^1(\mathbb{R})$. Then g is asymptotically null.

Proof. We argue by contradiction, assuming, for example, that $g(x)$ does not go to zero at $+\infty$. There then exists an $\epsilon > 0$ and a sequence of real numbers $\{a_n\}$, with $a_n \to \infty$ for $n \to \infty$, such that $|g(a_n)| > \epsilon$.

By taking an appropriate subsequence from $\{a_n\}$ if necessary, we can assume without loss of generality that the a_n are strictly increasing with n, that the $g(a_n)$ are all one sign (which we can take to be positive by considering $-g$ if necessary), and that for each n there is between a_n and a_{n+1} a point ξ_n at which $g(\xi_n) \leqslant \frac{1}{2}\epsilon$. The only assertion that needs any further explanation is the last one, and this becomes obvious when we recall that $g \in L_2(\mathbb{R})$, so that $g \geqslant \frac{1}{2}\epsilon$ can hold only on a set of finite meaure.

Let b_n be the first value of x less than a_n at which $g(x) = \frac{1}{2}\epsilon$, and let c_n be the first value greater than a_n at which $g(x) = \frac{1}{2}\epsilon$. These definitions are justified by the continuity of g. Write $I_n = (b_n, c_n)$. Then $\{I_n\}$ is a sequence of disjoint, non-empty and open intervals such that $g(\xi) \geqslant \frac{1}{2}\epsilon$ for $\xi \in I_n$, and $g(\xi) = \frac{1}{2}\epsilon$ at the end points of I_n. It follows that $m(I_n) = c_n - b_n$ must converge to zero since $g \in L_2(\mathbb{R})$.

Applying the mean-value theorem to g on I_n, we may infer the existence of points $\sigma_n \in (b_n, a_n)$ and $\tau_n \in (a_n, c_n)$ such that

$$g'(\sigma_n) = \frac{g(a_n) - g(b_n)}{a_n - b_n} \quad \text{and} \quad g'(\tau_n) = \frac{g(c_n) - g(a_n)}{c_n - a_n}.$$

Taking absolute values and estimating lower bounds for these quotients, we obtain

$$|g'(\sigma_n)| \geqslant \frac{\epsilon}{2(a_n - b_n)}, \quad |g'(\tau_n)| \geqslant \frac{\epsilon}{2(c_n - a_n)}.$$

One of the right-hand sides of these two inequalities must be at least $\epsilon/\{4(c_n - b_n)\}$. Thus a point $\lambda_n \in (b_n, c_n)$ is shown to exist such that

$$|g'(\lambda_n)| > \frac{\epsilon}{4(c_n - b_n)},$$

which, in light of the fact that $(c_n - b_n) \to 0$, contradicts the assumption that g' is bounded. Thus the assumption that $g(x)$ is not asymptotically null is shown to be incorrect, and the lemma is proven.

PROPOSITION. $g \in W_2^1(\mathbb{R}) \cap C^2(\mathbb{R})$ implies that g and g' are asymptotically null.

Proof. Both g and g' satisfy the condition of the lemma.

GLOSSARY OF SPECIAL SPACES

space	*defining property of elements*	*norm*
$\mathscr{C}_T$	$u(x,t)$ continuous on $\mathbb{R} \times [0, T]$	$\|u\|_{\mathscr{C}} = \sup_{x \in \mathbb{R},\, t \in [0,T]} |u(x,t)|$
$\mathscr{C}_T^{l,m}$	$\partial_x^i \partial_t^j u \in \mathscr{C}_T$ for $0 \leqslant i \leqslant l$, $0 \leqslant j \leqslant m$	$\sum_{i=0}^{l} \sum_{j=0}^{m} \|\partial_x^i \partial_t^j u\|_{\mathscr{C}}$
$\mathscr{L}_T$	$f(x,t) \in L_2(\mathbb{R})$ for each $t \in [0, T]$. f: $[0, T] \to L_2(\mathbb{R})$ is a continuous map.	$\|f\|_{\mathscr{L}} = \sup_{t \in [0,T]} \|f(x,t)\|_2$ $\left[\|u\|_2 = \left(\int_{-\infty}^{\infty} u^2 \mathrm{d}x\right)^{\frac{1}{2}}\right]$
$\mathscr{W}_T$	$u(x,t) \in W_2^1(\mathbb{R})$ for each $t \in [0, T]$. u: $[0, T] \to W_2^1(\mathbb{R})$ is a continuous map.	$\|u\|_{\mathscr{W}} = \sup_{t \in [0,T]} \|u(x,t)\|_{1,2}$ $[\|u\|_{1,2} = (\|u\|_2^2 + \|u_x\|_2^2)^{\frac{1}{2}} = E^{\frac{1}{2}}(u)]$
$\mathscr{G}$	$(g,f) \in \mathscr{G} \Rightarrow g \in W_2^1(\mathbb{R}) \cap C^2(\mathbb{R})$, $f \in \mathscr{L}_T \cap \mathscr{W}_T$	$\|(g,f)\|_{\mathscr{G}} = \|g\|_{1,2} + \|g\|_{C^2} + \|f\|_{\mathscr{L}} + \|f\|_{\mathscr{W}}$

References

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

Benjamin, T. B. 1967*b* Some developments in the theory of vortex breakdown. *J. Fluid Mech.* **28**, 65.

Benjamin, T. B. 1972 The stability of solitary waves. *Proc. R. Soc. Lond.* A (to appear).

Benjamin, T. B. & Barnard, B. J. S. 1964 A study of the motion of a cavity in rotating fluid. *J. Fluid Mech.* **19**, 193.

Broer, L. J. F. 1964 On the interaction of non-linearity and dispersion in wave propagation. I. Boussinesq's equation. *Appl. scient. Res.* B **11**, 273.

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.

Hadamard, J. 1923 *Lectures on Cauchy's problem in linear partial differential equations.* Yale University Press. (Dover reprint 1952.)

Hardy, G. H., Littlewood, J. E. & Polya, G. 1952 *Inequalities* (2nd ed.). Cambridge University Press.

Jeffreys, H. & Jeffreys, B. S. 1966 *Methods of mathematical physics* (3rd ed.). Cambridge University Press.

Kametaka, Y. 1969 Korteweg–de Vries equation. I. Global existence of smooth solutions. *Proc. Japan Acad.* **45**, 552.

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.* (v.), **39**, 422.

Lamb, H. 1932 *Hydrodynamics* (6th ed.). Cambridge University Press. (Dover reprint 1945.)

Lax, P. D. 1968 Integrals of nonlinear equations of evolution and solitary waves. *Communs pure appl. Math.* **21**, 467.

Leibovich, S. 1970 Weakly nonlinear waves in rotating fluids. *J. Fluid Mech.* **42**, 803.

Lions, J. L. 1969 *Quelques Méthodes de Résolution des Problèmes aux Limites Non Linéaires.* Paris: Dunod.

Meyer, R. E. 1967 Note on the undular jump. *J. Fluid Mech.* **28**, 209.

Miura, R. M. 1968 Korteweg–de Vries equation and generalizations. I. A remarkable explicit nonlinear transformation. *J. math. Phys.* **9**, 1202.

Miura, R. M., Gardner, C. S. & Kruskal, M. D. 1968 Korteweg–de Vries equation and generalizations. II. Existence of conservation laws and constants of motion. *J. math. Phys.* **9**, 1204.

Nirenberg, L. 1971 Pseudo-differential operators. *Proc. Symp. Pure Math., Global Analysis.* Amer. Math. Soc.

Peregrine, D. H. 1964 Calculations of the development of an undular bore. *J. Fluid Mech.* **25**, 321.

Pritchard, W. G. 1970 Solitary waves in rotating fluids. *J. Fluid Mech.* **42**, 61.

Rudin, W. 1966 *Real and complex analysis.* New York: McGraw-Hill.

Sjöberg, A. 1970 On the Korteweg–de Vries equation: existence and uniqueness. *J. math. Analysis Applic.* **29**, 569.

Smirnov, V. I. 1964 *A course of higher mathematics.* V. *Integration and functional analysis.* London: Pergamon.

Su, C. H. & Gardner, C. S. 1969 Korteweg–de Vries equation and generalizations. III. Derivation of the Korteweg–de Vries equation and Burgers equation. *J. math. Phys.* **10**, 536.

Temam, R. 1969 Sur un problème non linéaire. *J. Math. pures appl.* **48**, 159.

Whitham, G. B. 1967 Variational methods and applications to water waves. *Proc. R. Soc. Lond.* A **299**, 6.