J. Fluid Mech. (1966), *vol.* 25, *part* 2, *pp.* 241–270
Printed in Great Britain

Internal waves of finite amplitude and permanent form

By T. BROOKE BENJAMIN
Department of Applied Mathematics and Theoretical Physics,
University of Cambridge

(Received 6 September 1965)

A theory is derived for the class of long two-dimensional waves, comprising solitary and periodic cnoidal waves, that can propagate with unchanging form in heterogeneous fluids. The treatment is generalized to the extent that the waves are supposed to arise on a horizontal stream of incompressible fluid whose density and velocity are arbitrary functions of height, and the upper surface of the fluid is allowed either to be free or to be fixed in a horizontal plane. Explicit formulae for the wave properties and a general interpretation of the physical conditions for the occurrence of the waves are achieved without need to specify particular physical models; but in a later part of the paper, §4, the results are applied to three examples that have been worked out by other means and so provide checks on the present theory. These general results are also shown to accord nicely with the principle of 'conjugate-flow pairs' which was explained by Benjamin (1962*b*) with reference to swirling flows along cylindrical ducts, but which is known to apply equally well to flow systems of the kind in question here.

The theory reveals certain physical peculiarities of a type of flow model often used in theoretical studies of internal-wave phenomena, being specified so as to make the equation for the stream-function linear. In an appendix, some observations are also made regarding the 'Boussinesq approximation', which too is often used as a simplifying assumption in this field. It is shown, adding to a recent discussion by Long (1965), that finite internal waves may depend crucially on small effects neglected in this approximation.

1. Introduction

In order to define the class of internal waves that is the subject of this paper, it is helpful to refer first to the problem of long gravity waves upon homogeneous liquids. A satisfactory definition for the present purpose is suggested by recalling how the classical solitary- and cnoidal-wave solutions—as given for instance in Lamb's text-book (1932, §§252, 253)—are explained in relation to other chapters of water-wave theory, particularly to the non-linear 'shallow-water' theory due to Airy (Lamb, §§187, 188), which appeared originally to demonstrate the impossibility of continuous long waves having both finite amplitude and permanent shape. The essential explanation proceeds as follows (cf. Ursell 1953; Benjamin & Lighthill 1954; Long 1964).

Taking the depth of the undisturbed fluid as the unit of length, one is presented with two dimensionless scale parameters for consideration in developing approximate theories: the amplitude ϵ of the vertical displacements, and the minimum horizontal length l typical of the wave motion. Being meant to apply to very long waves, Airy's theory neglects the effects of vertical accelerations, which are a fraction $O(l^{-2})$ of the effects included. Hence it shows that waves of elevation steepen ahead of their crests at a rate proportional to ϵ times the velocity $c_0 = \sqrt{g}$ of infinitesimal waves with extreme length (Lamb, p. 279). On the other hand, the linearized 'surface-wave' theory (Lamb, ch. 9), retaining only the first term of an expansion in terms of ϵ/l, shows that a non-sinusoidal wave suffers dispersion of its steepest parts at a rate depending on the difference between the phase and group velocities of its predominant spectral components, which for long waves is found to become proportional to c_0/l^2 (see Jeffreys & Jeffreys 1946, §17.09). Considered together, the two approximate theories point to the possibility of an intermediate class of waves for which a cancellation of these two opposing tendencies may permit permanency of form. This important central ground was first covered by Korteweg & de Vries (1895), whose analysis took both ϵ and l^{-2} to be small but gave equal standing to these two parameters, hence proceeding to a first approximation for the effects of both finite amplitude and vertical accelerations. The class of steady waves forthcoming from this type of analysis comprises the solitary wave, for which the ratio ϵl^2 of the two small parameters has a maximum $O(1)$, cnoidal waves, for which ϵl^2 takes somewhat smaller values, and sinusoidal waves, which are given in the limit as ϵl^2 tends to zero.

Now, with regard to internal waves a third basic parameter comes into consideration, namely the scale for the density variations with height. This can be taken as b^{-1} if b is the fractional change in density over the total (unit) depth of fluid, and for the moment let us restrict the discussion to the case of waves propagating into a region of fluid initially at rest, so that b^{-1} is the only scale height for the undisturbed state. (We shall see later that, when there is a non-uniform horizontal flow initially, its vertical scale as well as b^{-1} may be important in the following respects.) An infinite number of internal-wave modes is generally possible in a given heterogeneous-fluid system, but for clearness we now fix attention on the first of them, which is the fastest moving, and denote by c_1 the speed of infinitesimal, extremely long waves in this mode. It is well known that $c_1 = O(bg)^{\frac{1}{2}}$, and since $b \ll 1$ in most practical applications c_1 may be considerably smaller than the speed of long 'surface waves' on fluid of the same depth. From the extensive body of theory available for internal waves of infinitesimal amplitude ϵ (e.g. see Lamb 1932, §235; Yih 1960*a*; Yanowitch 1962), it appears generally true that the speed of sinusoidal waves in a particular mode is a maximum for infinite wavelength. Hence a general conclusion may readily be derived similar to the one mentioned above regarding the dispersion of long non-sinusoidal waves: for such a wave in the first mode, the rate of dispersion is seen to be $O(c_1/l^2)$ [for an instance of the sort of simple argument that is required, see Benjamin & Barnard (1964, pp. 204–206)].

Unfortunately, no general 'shallow-water' theory for internal waves has yet been published to which we might have appealed for the other element in the proposed explanation. However, the required general result may be deduced readily enough from first principles. We shall not delay here to present a proper derivation, which could by itself fill a whole paper, and it must suffice to describe one possible argument in barest outline. This adapts a method given by Lamb (1932, § 175) for deriving certain results in Airy's theory, and considers a long internal wave of finite amplitude travelling in one direction only. Each vertical section through the wave is viewed in a frame of reference such that the motion relative to the undisturbed fluid far ahead is approximately steady, and conditions of mass and energy conservation are applied. Then the method of 'conjugate-flow' analysis explained by Benjamin (1962*b*) is used to find the local wave speed. It thus appears in general that this speed exceeds the infinitesimal-wave speed c_1 by a fraction $O(ba)$, where a is the local amplitude, so that as a whole the wave steepens—i.e. decreases its length scale l—at a rate $O(c_1 b\epsilon)$. [This conclusion is borne out by the results of recent work by Benney (1966), which is mentioned again below and is to be published in the near future.]

Hence, retracing the steps explaining the Korteweg & de Vries theory, we deduce that internal solitary and cnoidal waves are, at least in the first mode, characterized by scales such that $b\epsilon l^2 = O(1)$, although both $b\epsilon$ and l^{-2} may be arbitrarily small. The framework of this deduction clarifies both the physical and analytical basis of the present theory, which is accordingly identified as a counterpart in all essentials to the Korteweg & de Vries theory for long waves of permanent form. However, in formally developing the approximation that gives solitary and cnoidal waves, we shall conveniently take just ϵ rather than $b\epsilon$ as the expansion parameter and assume $l = O(\epsilon^{-\frac{1}{2}})$; in other words, we shall proceed as if $b = O(1)$, taking no advantage of the usual property that $b \ll 1$, and allow the general, crucial dependence of the solutions on b to emerge as an end-result of the analysis.

The theory of internal solitary waves was initiated by Keulegan (1953) and Long (1956) who investigated two-fluid systems with fixed upper boundaries. The corresponding problem when the upper boundary is free was treated by Peters & Stoker (1960), and they also examined at length the more difficult problem of solitary waves in a fluid whose density decreases exponentially with height. It was recognized by them that their theory would yield periodic solutions, i.e. cnoidal waves, in the same category of approximation as the solitary-wave solution; but they did not attempt to interpret the circumstantial distinctions between the possible solutions in this general class. The method of analysis used by Peters & Stoker has been extended by Shen (1964, 1965) to deal with solitary waves in compressible heavy fluids, and very recently Benney (1966) has developed a new general method—certainly no less successful than the present one in so far as the results of the two can be compared—for treating unsteady as well as steady non-linear waves in stratified shear flows.

A very powerful analysis of some aspects of the present problem has been made by Ter-Krikorov (1963), using methods that he had previously applied with great effectiveness to the study of solitary waves on a stream of homogeneous fluid with vorticity (a problem treated in a somewhat different way by Benjamin 1962*a*). Allowing wide generality in the specification of primary density and velocity distributions, he succeeded in the difficult task of proving the convergence of approximation schemes that yield solitary and cnoidal waves, thus establishing the existence of waves with permanent form in the present type of system. This achievement is allied to the existence proof that Friedrichs & Hyers (1954) found for the classical solitary wave, which was extended to cnoidal waves by Littman (1957).

Taking a line of approach complementary to Ter-Krikorov's fundamental work on the analytical problem, the present paper aims particularly to elucidate the practical applications of the theory. The phenomena of lee waves and internal bores are brought into the discussion, for instance, and the object is to fit them into a general physical interpretation. In fact the accomplishment considered to be of most value in the paper is to build up a classification of various internal-wave phenomena that is a precise parallel to the one achieved in Benjamin & Lighthill's (1954) presentation of classical cnoidal-wave theory, which was shown to account neatly for the undular hydraulic jump and several other well-known effects in open-channel flows (see also Benjamin 1956). A note will also be made of certain connexions with the author's theory of vortex breakdown (1962*b*, 1965), which virtually proved the possibility of waves of the present type in swirling flows along cylindrical ducts but did not demonstrate them explicitly.

2. Formulation of the problem and basic equations

The primary state upon which waves are to be superposed consists of a layer of frictionless fluid, of depth h_0, resting on a rigid horizontal bottom. Its upper surface may be either free, i.e. at pressure $p = 0$, or fixed by a rigid horizontal plane. The density ρ of the fluid is an arbitrary function of height y above the bottom, though necessarily decreasing upwards to ensure 'static stability'. There may be a primary parallel flow, the horizontal velocity being an arbitrary function $U(y)$. The dynamic stability of such a system to small disturbances is a difficult question which will not be taken up here, but reference can be made to papers by Miles (1961, 1963) and Howard (1961) for an account of modern developments in the subject. A sufficient condition for stability, which will be assumed here to clear this aspect, is that $-g\rho^{-1}d\rho/dy \geqslant \frac{1}{4}(dU/dy)^2$ (see Howard 1961). Presumably, by taking this condition to be rather more than marginally satisfied, we can also ensure the stability of the steady wave disturbances having small but finite amplitude that are to be investigated below.

We take a horizontal axis x with respect to which the waves to be analysed are stationary. Thus the respective primary velocity in the x-direction becomes $W = c+U$, if c be the wave velocity *upstream* in a fixed frame of reference. The special cases of propagation into a fluid at rest ($U = 0$) and of downstream propagation ($c+U < 0$) are obviously included in this general representation, and for the most part there is no need to distinguish between them in the development of the theory (cf. Benjamin 1962*a*, p. 101). In §2 of the author's previous paper just cited several physical considerations were discussed relating to flow models of the present kind, in particular with regard to the implications of an inviscid initial flow with vorticity, and this reference may serve in lieu of a basic physical discussion here.

Let $\psi(x,y)$ denote the stream-function for a steady two-dimensional flow in the specified frame of reference, so that the horizontal and vertical components of velocity are given by

$$u = \psi_y, \quad v = -\psi_x. \tag{2.1}$$

It is assumed that no reversal of flow occurs in any vertical section, which means that ψ varies monotonically with y. On the further assumption that the fluid is incompressible and non-diffusive, density is constant along each streamline, and this property is representable by

$$\rho = \rho(\psi). \tag{2.2}$$

The dynamical condition on each streamline is that the stagnation pressure or 'total head' is constant; thus

$$H = p + \tfrac{1}{2}\rho(u^2+v^2) + g\rho y = H(\psi). \tag{2.3}$$

Consider next the 'flow force', equal to the horizontal pressure force plus the flux of horizontal momentum, for a surface of unit span extending between any two points A and B in the flow plane: that is

$$\mathscr{S} = \int_A^B (p\,dy + \rho u\,d\psi) = \int_A^B \{(p+\rho\psi_y^2)\,dy + \rho\psi_y\psi_x dx\}.$$

By use of (2.3) to eliminate p, this becomes

$$\mathscr{S} = \int_A^B [\{H + \tfrac{1}{2}\rho(\psi_y^2 - \psi_x^2) - g\rho y\}\,dy + \rho\psi_y\psi_x dx]. \tag{2.4}$$

In particular, for a vertical section ($dx = 0$) between the bottom $y = 0$ and the upper boundary $y = h$, let us use a new symbol S for the flow force; thus, for this special case,

$$\mathscr{S} = S = \int_0^h \{H + \tfrac{1}{2}\rho(\psi_y^2 - \psi_x^2) - g\rho y\}\,dy. \tag{2.5}$$

This quantity has a role of central importance in the subsequent analysis. Note that $h = h_0$ (const.) when the upper boundary is fixed, but $h = h(x)$ when it is free and so can be displaced by the wave motion.

2.1. *A new derivation of Long's equation*

From the general flow-force integral (2.4) certain well-known results can be derived in a simple way, which seems worth some discussion as a novel alternative to the usual, more direct methods of derivation. We first note that the value of $\mathscr{S}$ given by (2.4) is independent of the path between fixed points A and B in a simply-connected region of the flow, because in the absence of any horizontal external force the flow force across a closed surface (e.g. one demarked by two alternative paths between A and B) must be zero. We now consider the formal variation in $\mathscr{S}$ according to (2.4) when the path between A and B is changed from one described by the equation $x = \xi(y)$ to one described by $x = \xi(y)+\delta\xi(y)$. In evaluating the variation of the integral, keeping the element dy fixed, we may put

$$\begin{aligned}\delta(\tfrac{1}{2}\psi_x^2) &= \psi_x\psi_{xx}\delta\xi,\\ \delta(\tfrac{1}{2}\psi_y^2) &= \psi_y\psi_{xy}\delta\xi,\\ \delta(dx) &= (\delta\xi)_y\,dy,\\ \delta H, \delta\rho &= (H', \rho')\,\psi_x\delta\xi,\end{aligned}$$

where H' and ρ' denote $dH/d\psi$ and $d\rho/d\psi$. Hence the first variation, which must of course vanish, is given by

$$\delta\mathscr{S} = \int_A^B [\{H' + \tfrac{1}{2}\rho'(\psi_y^2 - \psi_x^2) - gy\rho' - \rho\psi_{xx}\}\,\psi_x\delta\xi + \rho\psi_y\psi_{xy}\delta\xi + \rho\psi_x\psi_y(\delta\xi)_y]\,dy.$$

Integrating the final term in the integral by parts and using the condition that $\delta\xi = 0$ at the end-points, we obtain the result

$$\delta\mathscr{S} = 0 = \int_A^B \psi_x\mathscr{F}\,\delta\xi\,dy, \tag{2.6}$$

with

$$\mathscr{F} = \rho\nabla^2\psi + \rho'\{\tfrac{1}{2}(\psi_x^2+\psi_y^2)+gy\} - H'. \tag{2.7}$$

Since $\delta\xi$ is arbitrary and ψ_x cannot be zero everywhere except in a strictly horizontal parallel flow, it follows that

$$\mathscr{F} = 0. \tag{2.8}$$

This second-order partial differential equation for the stream-function, which is generally non-linear, was derived by Dubreil-Jacotin (1937) and by Long (1953) directly from the equations of motion, and it has been used by them and many others as the starting-point for studies of steady internal waves. It is commonly called Long's equation.

Note, incidentally, that an equation akin to (2.6) but not as informative may be obtained by differentiation of (2.5) with respect to x. After the differentiation and an integration by parts, the terms outside the integral sign are seen to vanish in consequence of the boundary condition at $y = h$, either $p = 0$ or $dh/dx = 0$ for a free or fixed surface respectively. The result is

$$\frac{dS}{dx} = \int_0^h \psi_x\mathscr{F}\,dy. \tag{2.9}$$

We have $dS/dx = 0$ as an obvious physical property in the absence of horizontal external forces, and so (2.9) is consistent with (2.8). But (2.8) is not proved by the present result, since ψ_x is not an arbitrary function of y.

2.2. *Yih's transformation and the identification of linear systems*

A transformation introduced by Yih (1960*b*; see also Long 1953, p. 46) which simplifies Long's equation is now recalled, revealing a point of great significance to our subsequent analysis. Because of the property (2.2), there exists a pseudo-stream-function $\tilde{\psi}(x, y)$ such that

$$\rho^{\frac{1}{2}}u = \tilde{\psi}_y, \quad \rho^{\frac{1}{2}}v = -\tilde{\psi}_x. \tag{2.10}$$

In terms of this, (2.4) may be rewritten

$$\mathscr{S} = \int_A^B [\{H(\tilde{\psi}) + \tfrac{1}{2}(\tilde{\psi}_y - \tilde{\psi}_x) - gy\rho(\tilde{\psi})\}\, dy + \psi_x \psi_y\, dx], \tag{2.11}$$

and (2.5) rewritten

$$S = \int_0^h \{H(\tilde{\psi}) + \tfrac{1}{2}(\tilde{\psi}_y^2 - \tilde{\psi}_x^2) - gy\rho(\tilde{\psi})\}\, dy. \tag{2.12}$$

By means of the variational argument used to obtain (2.8), it is readily found from (2.11) that

$$\nabla^2 \tilde{\psi} + gy\frac{d\rho}{d\tilde{\psi}} - \frac{dH}{d\tilde{\psi}} = 0 \tag{2.13}$$

(cf. Yih 1960*b*, equation (10); Ter-Krikorov 1963, equation (4)). This equation may also be derived, of course, by substitution of (2.10) into (2.8).

Exact solutions of (2.13), and *a fortiori* of (2.8), have so far been obtained only in special cases where the equation is *linear*. The chief aim of Yih's paper (1960*b*) was to identify the complete class of such cases, and his simpler equation (2.13) is clearly better suited to this purpose than Long's equation (2.8). The necessary and sufficient condition for (2.13) to have a linear form is simply that both ρ and H are polynomials of degree no higher than quadratic in $\tilde{\psi}$. It follows at once from this that, if the stream-function is expressed as the sum of a primary function and a perturbation, the expansion of the flow-force integral (2.12) in terms of the perturbation will *terminate at second order* when the upper boundary is fixed. This analytical property will be shown later to imply certain physical peculiarities of heterogeneous-fluid systems specified so as to make the equation for the stream-function linear. Such systems have frequently been taken as tractable models for lee-wave and kindred phenomena (e.g. see Long 1958), and it has generally been assumed that no qualitative difference exists between them and the much wider class of systems for which (2.13) is non-linear. The distinctions that will be pointed out presently do not in any sense undermine previous theoretical work on the basis in question, but they do provide some much-needed clarification of this extensive branch of the subject.

3. Main analysis

The essence of our method is to regard the definition of the flow force S as an integral equation and, in effect, to find the solution by successive approximations. It must be anticipated from the start, however, that solutions with the same characteristics as ordinary solitary and cnoidal waves will not emerge until an advanced stage in the process of approximation, in fact at third order in the expansion parameter. Precisely the same type of singular-perturbation problem is presented as in classical solitary-wave theory (see Ursell 1953), and as in that context no reliable qualitative conclusion can be drawn before the required approximation is complete. In the case of a fixed upper boundary, equation (2.5) or (2.12) would make a suitable starting point, but an analysis proceeding directly therefrom would encounter considerable difficulty in dealing with the case of a free boundary. To cover both cases with equal simplicity, the following device is used (cf. Benjamin 1962*a*, p. 103).

3.1. *Transformation of variables*

Take the height y of the individual streamlines as the dependent variable, and let the independent variables be x and the height, η say, of the respective streamlines in the original parallel flow; thus

$$y = y(x, \eta). \tag{3.1}$$

This representation is unique on the assumption that no reversal of flow occurs anywhere; i.e. all streamlines connect without bifurcation to every vertical section and cross it in the same order of height. The idea is illustrated in figure 1.

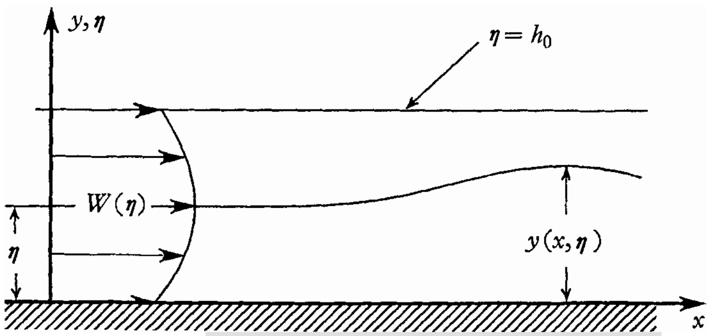


FIGURE 1. Definition sketch showing primary velocity profile $W(\eta)$. The height y of the streamlines is considered as a function of their height η in the primary flow and of the horizontal co-ordinate x.

Since $d\psi/d\eta = W(\eta)$ by definition, we have

$$u = W(\eta)\frac{1}{y_\eta}, \quad v = W(\eta)\frac{y_x}{y_\eta}. \tag{3.2}$$

Also, ρ and H are expressible as functions of η alone, since ψ is uniquely determined by η. Hence, writing $\rho W^2 = Q(\eta)$, we get from (2.5)

$$S = \int_0^{h_0} \left\{H + \tfrac{1}{2}Q\left(\frac{1-y_x^2}{y_\eta^2}\right) - g\rho y\right\} y_\eta \, d\eta, \tag{3.3}$$

and for the value of S in the primary flow, where $y = \eta$,

$$S_0 = \int_0^{h_0} \{H + \tfrac{1}{2}Q - g\rho\eta\}\, d\eta. \tag{3.4}$$

Note that the integral (3.3) has the fixed limit h_0 even if the upper boundary is free, in contrast to the integral (2.5). The advantages of the transformation (3.1) owe essentially to this feature.

3.2. *Method of successive approximation*

To obtain approximate results for the vertical displacements of the streamlines from their original level, we put

$$y = \eta + \epsilon\zeta(x, \eta), \tag{3.5}$$

regarding ϵ as a small number, and expand $S-S_0$ in powers of ϵ.† It can be expected that the coefficient of ϵ^1 in the expansion will vanish identically, in view of the general property of all wave systems that flow-force perturbations (i.e. 'wave resistance') are no larger than of second order in displacements (cf. Lamb 1932, §249; Benjamin 1962*b*, Appendix, §(*c*)). The demonstration of this is quite interesting in the present case, however, and the details are worth noting.

To $O(\epsilon)$, we get from (3.3) and (3.4)

$$\begin{aligned} S-S_0 &= \epsilon\int_0^{h_0}\{(H-\tfrac{1}{2}Q-g\rho\eta)\,\zeta_\eta - g\rho\zeta\}\,d\eta \\ &= \epsilon\int_0^{h_0}(p_0\zeta_\eta - g\rho\zeta)\,d\eta. \end{aligned} \tag{3.6}$$

Here p_0 denotes the pressure in the primary flow. But p_0 satisfies the hydrostatic law $dp_0/d\eta = -g\rho$. Hence (3.6) gives

$$S-S_0 = \epsilon[p_0\zeta]_0^{h_0}.$$

These integrated terms vanish because $\zeta = 0$ on the bottom $\eta = 0$, and at the upper boundary $\eta = h_0$ either $\zeta = 0$ (fixed surface) or $p_0 = 0$ (free surface).

3.3. *Infinitesimal disturbances*

The well-established theory of stable infinitesimal waves in heterogeneous liquids (e.g. see Yih 1960*a*; Yanowitch 1962) is recovered when the expansion of $S-S_0$ is taken to second order. We first require the linearized form of the boundary condition for a free surface, adding to the exact conditions $\zeta = 0$ at $\eta = 0$ and, for a fixed upper boundary, $\zeta = 0$ at $\eta = h_0$. Evaluating the total head at the free surface, first for the primary flow and then for the disturbed flow, we have

$$H(h_0) = \tfrac{1}{2}Q + g\rho h_0 = \tfrac{1}{2}Q\left(\frac{1+y_x^2}{y_\eta^2}\right) + g\rho y. \tag{3.7}$$

Linearization in ϵ after substitution of (3.5) gives directly

$$Q\zeta_\eta = g\rho\zeta \quad \text{at} \quad \eta = h_0. \tag{3.8}$$

To $O(\epsilon^2)$, we get from (3.3) and (3.4)

$$\begin{aligned} S-S_0 &= \epsilon^2\int_0^{h_0}\{\tfrac{1}{2}Q(\zeta_\eta^2-\zeta_x^2) - g\rho\zeta\zeta_\eta\}\,d\eta \\ &= \tfrac{1}{2}\epsilon^2\left\{[\zeta(Q\zeta_\eta - g\rho\zeta)]_0^{h_0} - \int_0^{h_0}\{Q\zeta_x^2 + (Q\zeta_\eta)_\eta\zeta - g\rho_\eta\zeta^2\}\,d\eta\right\}. \end{aligned}$$

The integrated terms vanish in consequence of the boundary conditions, with either (3.8) or $\zeta = 0$ at $\eta = h_0$. Hence

$$S-S_0 = -\tfrac{1}{2}\epsilon^2\int_0^{h_0}\{Q\zeta_x^2 + (Q\zeta_\eta)_\eta\zeta - g\rho_\eta\zeta^2\}\,d\eta. \tag{3.9}$$

† It is implied here that ϵ represents the typical magnitude of $y-\eta$ when expressed in units of the original total depth h_0; thus $\zeta = O(1)$ in these units.

To the same order of approximation, and again making use of the boundary conditions, we find that

$$\frac{d}{dx}(S-S_0) = -\epsilon^2\int_0^{h_0}\zeta_x\{Q\zeta_{xx}+(Q\zeta_\eta)_\eta-g\rho_\eta\zeta\}\,d\eta.$$

Since $d(S-S_0)/dx = 0$ in the absence of horizontal external forces, this equation suggests that

$$Q\zeta_{xx}+(Q\zeta_\eta)_\eta-g\rho_\eta\zeta = 0, \tag{3.10}$$

and comparison with (2.9) confirms that (3.10) is equivalent to the linearized form of Long's equation. Equation (3.10) can be obtained in several other ways, for instance by the variational method used in §2 or by direct linearization of Long's equation after the transformation (3.1) is introduced; but a more detailed proof need not delay us. This is, clearly enough, the equation to be satisfied by infinitesimal steady disturbances, and it may easily be shown to correspond to the other forms of basic linearized equation that have been used by previous authors (e.g. Lamb 1932, §235; Yih 1960*a*; Yanowitch 1962).

Substitution of (3.10) into (3.9) gives

$$S-S_0 = -\tfrac{1}{2}\epsilon^2\int_0^{h_0}Q\{\zeta_x^2-\zeta\zeta_{xx}\}\,d\eta, \tag{3.11}$$

a neat equation with several points of physical interest (cf. Benjamin 1962*b*, equation (A 22)). First, it proves that a flow-force reduction (i.e. positive wave resistance) is a necessary property of *sinusoidal* internal waves produced without energy loss, just as is well known to be the case for gravity waves on a homogeneous liquid. Note, however, that the wave resistance S_0-S becomes progressively smaller, for a given amplitude factor ϵ, as the wavelength increases and hence the x-derivatives in (3.11) diminish in magnitude. This is readily appreciated to be an essential property of infinitesimal waves in any system that, as will be shown later for the present one, can manifest finite waves of the solitary and cnoidal type. Equation (3.11) also shows that $S = S_0$ for an infinitesimal disturbance having real exponential dependence on x. This case includes the outskirts of a solitary wave, where the displacements caused by the wave motion become very small and hence the linearized theory becomes applicable (see Benjamin 1962*b*, pp. 596, 600).

3.4. *The eigenvalue problem for sinusoidal infinitesimal waves*

Putting $\zeta(x,\eta) = \phi(\eta)\sin(\alpha x+\nu)$, we get from (3.10) and the boundary conditions the alternative systems

$$\frac{d}{d\eta}\left(Q\frac{d\phi}{d\eta}\right)-\left(Q\alpha^2+g\frac{d\rho}{d\eta}\right)\phi = 0, \tag{3.12}$$

$$\left.\begin{array}{l}\phi = 0 \quad \text{at} \quad \eta = 0,\\ \text{either}\quad \phi = 0\ \text{(fixed)}\quad \text{or}\quad Q\dfrac{d\phi}{d\eta} = g\rho\phi\ \text{(free)}\quad \text{at}\quad \eta = h_0.\end{array}\right\} \tag{3.13}$$

First suppose that $Q = \rho(c+U)^2$ is specified. Then, for either form of the upper boundary condition, the problem of finding values of α^2 for which (3.12) and (3.13) have non-trivial solutions is a standard Sturm–Liouville problem. According to a well-known theorem (Ince 1926, § 10.61), there exists an infinite set of real eigenvalues $\alpha_0^2 > \alpha_1^2 > \alpha_2^2 > \ldots$ having no limit-point except $\alpha^2 = -\infty$. Only a limited number, if any, of these can be positive, which means that only this number of sine-wave forms of ζ is possible for a given U and wave velocity c. As will be explained presently, a given primary state of flow relative to x, having prescribed velocity $c+U$ in this reference frame, may be described as subcritical or supercritical with regard to a certain wave-mode accordingly as the respective eigenvalue α^2 is positive or negative.

The alternative problem of finding the wave velocity for a given α^2 is of the standard Sturm–Liouville type only when $U = \text{const.}$ and the upper surface is fixed. Then it readily appears that, for every real α^2, there exists an infinite sequence of positive eigenvalues $gh_0/(c+U)^2$ whose only limit-point is $+\infty$ (Yih 1960*a*). The problem is not of the standard type when the upper surface is free, since then c occurs in one of the boundary conditions as well as in the differential equation. However, Yanowitch (1962) has solved the modified problem and has proved the existence of a similar infinite sequence of eigenvalues in the case $U = 0$, from which follows an obvious generalization to the case $U = \text{const.}$ A more serious difficulty is presented when a non-uniform U is specified: certain bounds are then predetermined for c since, as may easily be appreciated as a physical condition (see Benjamin 1962*a*, p. 100), the function Q cannot be allowed to vanish anywhere in $(0, h_0)$ if there is to be a steady wave motion without any reversal of flow (i.e. subject to the assumption explained below (3.1)). In this case, therefore, an admissible value of c is not necessarily forthcoming from either of the systems (3.12), (3.13) for arbitrary α^2. Disregarding this possible limitation for the moment, we may proceed by means of Picone's comparison theorem (Ince 1926, § 10.31) to make an important general deduction about the dependency of c upon α^2. Applied to the present systems, the theorem shows that, for either form of the upper boundary condition, a reduction in α^2 must always be accompanied by an increase in Q; thus the wave speed $|c|$ always increases with wavelength. This suggests that certain wave speeds can generally be found for which one of the possible values of α^2 vanishes, a result established definitely by Yih's and Yanowitch's analyses in the case $U = \text{const.}$

Presuming the general truth of this possibility, we define a *critical* state of flow to occur when $Q \to Q_n$ $(n = 0, 1, 2, \ldots)$ such as to make $\alpha_n^2 \to 0$. The corresponding eigenfunction is denoted by $\phi_n(\eta)$, and by definition it satisfies

$$\frac{d}{d\eta}\left(Q_n \frac{d\phi_n}{d\eta}\right) - g\frac{d\rho}{d\eta}\phi_n = 0 \tag{3.14}$$

together with the boundary conditions (3.13). In physical terms, a critical flow is one that can support infinitesimal stationary waves of indefinitely great length; and, since

$$Q_n = \rho(c_n + U)^2, \tag{3.15}$$

the Q_n are determined by the long-wave velocities c_n with respect to a fixed frame of reference. By Sturm's fundamental comparison theorem (Ince, § 10.3), the rate of oscillation of the solution of (3.14) increases with decreasing Q_n;

hence we may order the possible solutions according to the property that ϕ_{n+1} oscillates just once more than ϕ_n in the interval $(0, h_0)$, implying that

$$Q_0 > Q_1 > Q_2 > \ldots$$

when the Q_n are given by (3.15) with the same function U on the right-hand side.

In the special case $U = \text{const.}$, we know from the findings by Yih and Yanowitch mentioned above that we can find an infinite number of critical states represented by different Q_n. However, it is $U + c_n$ rather than c_n that is determined as an eigenvalue in this case, and so with regard to particular physical problems the number of *relevant* critical states may be limited. For instance, if U is a positive constant and only upstream propagation ($c > 0$) is in question, then only those eigenvalues $U + c_n$ that exceeded the given value of U are relevant—which may be none if U is large enough. This easy example suggests how we should interpret the foregoing definition of critical states when $U(\eta)$ is not a constant. Without inquiring into the generality of the definition for an arbitrary function U, we can merely consider its *ad hoc* use in particular problems. Thus upstream propagation might be specifically in view, for instance, and for the prescribed U the definition would give a limited number, if any, of critical states with $c > 0$ as required. Nevertheless, it appears likely that, when *downstream* propagation is specified, an infinite number of relevant eigensolutions always exists. Though a formal proof of this possibility will not be attempted here, we may note that, by a choice of $-c$ greater than the maximum of U, Q can be made arbitrarily small in part of the interval $(0, h_0)$ while remaining positive throughout; hence the rate of variation for the solution of (3.12) or (3.14) can be increased indefinitely, and from this the result in question may be inferred (cf. Benjamin 1962*b*, footnote on p. 627).

We reserve the classification $n = 0$ for the first wave-mode when the upper boundary is *free*. It is readily seen that ϕ_0 does not oscillate in $(0, h_0)$, and the respective solution of (3.12) represents the shortest possible sine wave for a given phase velocity c. This mode is similar to a surface wave on a homogeneous fluid, and the first true internal wave corresponds to $n = 1$.

With respect to the nth mode, the condition $|c| > |c_n|$, hence $Q > Q_n$, is termed *supercritical*. Comparison between (3.12) and (3.14) by means of Picone's theorem shows that, for either form of the upper boundary condition, the respective eigenvalue α^2 (i.e. for a ϕ with the same number of oscillations in $(0, h_0)$ as ϕ_n) is necessarily negative, which means that sinusoidal stationary waves are impossible. For a *subcritical* state, with $|c| < |c_n|$ and so $Q < Q_n$, the eigenvalue α^2 is positive and therefore sinusoidal waves in the particular mode can occur. This division of flow states is essentially the same as can be made for open-channel flows of homogeneous fluids and for swirling flows along ducts of circular cross-section (see Benjamin 1962*b*, § 3).

3.5. *Non-linear theory for long waves*

An approximation having the same status as the first-order theory of ordinary solitary and cnoidal waves is forthcoming when the flow-force parameter $S - S_0$ is evaluated to $O(\epsilon^3)$ from the integrals (3.3) and (3.4). To establish this approximation the following properties are assumed by analogy with well-known results from classical solitary-wave theory, being confirmed *a posteriori* by the present results (see remarks in § 1).

(i) The horizontal scale of the relevant wave motions is $O(\epsilon^{-\frac{1}{2}})$. Accordingly, a compressed measure of horizontal distance is defined by $X = \epsilon^{\frac{1}{2}}x$, so that

$$\zeta_x^2 = \epsilon\zeta_X^2, \tag{3.16}$$

where ζ_X^2 is $O(1)$.

(ii) The departure from a critical condition of flow (e.g. the difference between the actual wave speed c and c_n) is $O(\epsilon)$. Hence we put

$$Q = Q_n + \epsilon\gamma_n, \tag{3.17}$$

where $\gamma_n > 0$ for supercritical and $\gamma_n < 0$ for subcritical conditions.

(iii) The wave resistance $S_0 - S$ is $O(\epsilon^3)$. This assumption is, of course, the explicit basis of the proposed procedure for deriving the non-linear theory, and it now appears to be consistent with the result (3.11) of the infinitesimal-disturbance theory when coupled with the assumption (i). We accordingly write

$$S_0 - S = \epsilon^3 s. \tag{3.18}$$

(iv) Although so far no energy loss has been allowed between the primary and disturbed flows, so that H remains the same function of η everywhere, it will be a great advantage with regard to eventual physical interpretations for us to recognize at this point that the theory can account for an energy loss that is $O(\epsilon^3)$. If we write

$$\int_0^{h_0} (H_0 - H)\, d\eta = \epsilon^3 r \tag{3.19}$$

and assume $r = O(1)$, this term is simply added to the third-order approximation derived from (3.3) and (3.4), while the other terms are the same as when derived under the assumption $r = 0$.†

Introducing (3.16), (3.18) and (3.19), we get from (3.3) and (3.4)

$$\begin{aligned} 2\epsilon^3(r-s) &= \int_0^{h_0} \{\epsilon^2(Q\zeta_\eta^2 - 2g\rho\zeta\zeta_\eta) - \epsilon^3 Q(\zeta_X^2 + \zeta_\eta^3)\}\, d\eta \\ &= \epsilon^2[\zeta(Q\zeta_\eta - g\rho\zeta)]_0^{h_0} - \int_0^{h_0} [\epsilon^2\{(Q\zeta_\eta)_\eta - g\rho\zeta\}\zeta + \epsilon^3 Q(\zeta_X^2 + \zeta_\eta^3)]\, d\eta. \end{aligned} \tag{3.20}$$

Introducing (3.17), putting

$$\zeta = f(X)\,\phi_n(\eta), \tag{3.21}$$

† It is perhaps worth further emphasis that this gross representation of possible dissipative processes is strictly sufficient for the deduction of wave properties, provided the energy loss envisaged is indeed only $O(\epsilon^3)$. A more refined definition of $\epsilon^3 r$ may be required, however, if the loss occurs through 'breaking' at the front of a wave-train, a case that will be discussed later. The dissipation of energy in turbulence then may be accompanied by some significant mixing of the fluid, i.e. by turbulent diffusion of the density gradients. Provided the effects of this are only $O(\epsilon^3)$, they too can by accommodated by the theory. For instance, mixing will tend to raise the centroid of the density distribution and so increase the negative value of the final term in (3.4). If we consider all perturbations to the integral (3.4) and define $S_0 - \epsilon^3 r$ to be its value for the 'ground state' that is produced by the dissipative process, and upon which the waves in question are to be formed, then we effectively cover the possible effects of mixing as well as the reduction in H.

and using the fact that ϕ_n satisfies (3.14) and the boundary conditions (3.13), we find that (3.20) is satisfied, all terms to $O(\epsilon^3)$ being cleared, if

$$2(r-s) = f^2\left[\gamma_n\phi_n\frac{d\phi_n}{d\eta}\right]_0^{h_0} - f^2\int_0^{h_0}\phi_n\frac{d}{d\eta}\left(\gamma_n\frac{d\phi_n}{d\eta}\right)d\eta$$
$$-f_X^2\int_0^{h_0}Q\phi_n^2\,d\eta - f^3\int_0^{h_0}Q_n\left(\frac{d\phi_n}{d\eta}\right)^3 d\eta.$$

[Note that Q is retained in the coefficient of f_X^2, rather than Q_n being substituted as would be consistent with the other approximations. Thus the length scale of the final solutions is determined with somewhat greater precision than is strictly necessary at this order of approximation. The justification for such an improvement was discussed by Benjamin (1962*a*, see note on pp. 109, 110) with regard to solitary waves in homogeneous fluids, and the same considerations apply here.] Upon integration by parts of the second term on the right-hand side, the integrated terms cancel. There follows the nicely symmetrical result

$$If_X^2 = Jf^2 - Kf^3 + 2(s-r), \tag{3.22}$$

in which

$$I = \int_0^{h_0} Q\phi_n^2\,d\eta, \tag{3.23}$$

$$J = \int_0^{h_0}\gamma_n\left(\frac{d\phi_n}{d\eta}\right)^2 d\eta, \tag{3.24}$$

$$K = \int_0^{h_0} Q_n\left(\frac{d\phi_n}{d\eta}\right)^3 d\eta. \tag{3.25}$$

Equation (3.22) determines the horizontal distribution of the wave elevation. Note that this equation has the same form whether the upper boundary is fixed or free; the boundary condition affects the final results only through the eigenfunction ϕ_n. Note also that the coefficient $I > 0$, and that $J > 0$ or $J < 0$ accordingly as the flow state is supercritical or subcritical: K may be either positive or negative and, as will be seen presently, its sign determines whether the waves are of elevation or depression. As must be true for consistency, it may easily be confirmed that the arbitrary constant multiplying ϕ_n as defined will not affect the result for ζ obtained from (3.21) and (3.22).

While being formally correct to $O(\epsilon^3)$, the present approximation is open to the following practical objection, to meet which a certain adjustment needs to be made. Taking h_0 as the unit of length, c_n as the unit of velocity, and say $\rho(0)$ as the unit of density, one finds that for $n = 0$ the coefficient K is $O(1)$, but for $n = 1$ it is $O(b)$, where b is the magnitude of the density variation or, more generally, of the variation of Q_1. The latter result may be seen from the expression (A 2) for K given in the Appendix, and it is borne out by Example 3 in § 4 below. Now, for our approximation to be reliable, presumably the third-order terms retained in the expansion of the flow-force integral (3.3) should greatly exceed the next term in the expansion with (3.21) substituted. Thus it appears that we require $K \gg \epsilon L$, where

$$L = \int_0^{h_0} Q_n\left(\frac{d\phi_n}{d\eta}\right)^4 d\eta.$$

But L is found to be $O(1)$ for all n, and so at first sight it seems that for internal waves with $n = 1$ the theory may be limited to amplitudes for which $\epsilon \ll b$. Since b is very small in most practical applications, this would be a severe restriction on the usefulness of the theory.

Fortunately, the limitation may be removed by the simple expedient of putting $\phi_n(y)$ instead of $\phi_n(\eta)$ in (3.21). That is, we take

$$\begin{aligned}\zeta &= f(X)\,\phi_n(y) \\ &= f(X)\,\phi_n(\eta+\epsilon\zeta) \\ &= f\{\phi(\eta)+\epsilon\zeta\phi_\eta(\eta)+\tfrac{1}{2}\epsilon^2\zeta^2\phi_{\eta\eta}(\eta)+\ldots\} \\ &= f\phi+\epsilon f^2\phi\phi_\eta+\epsilon^2 f^3(\tfrac{1}{2}\phi^2\phi_{\eta\eta}+\phi\phi_\eta^2)+\ldots\end{aligned} \tag{3.26}$$

and substitute in the expansion of $S-S_0$ as before. To third order in ϵ, an equation for $f(X)$ identical with (3.22) is obtained. But after lengthy calculations, which do not seem worth reproducing here, the fourth-order remainder is found to be the sum of various terms that are all $O(b\epsilon)$ or smaller relative to the terms retained in the third-order approximation. Thus now the only condition on the validity of the approximation appears to be that $b\epsilon$ should be small. Long (1965) has made a comparable analysis using y rather than η as an independent variable, though only considering the case of a fixed upper boundary, and the accuracy of his approximate results also appears to depend on this condition alone.

The expression (3.26) with f determined by (3.22) will therefore be adopted as the more accurate approximation for internal waves when b is small.

3.6. *The solitary wave*

The physical circumstances of the wave are that it arises from the primary flow without change of energy or flow force; thus $r = s = 0$. Under these conditions a non-trivial real solution of (3.22) exists only if $J > 0$ (supercritical case). The solution is

$$\left.\begin{aligned} f &= a\,\mathrm{sech}^2\,\kappa X, \\ \text{where}\quad a &= J/K, \quad \kappa = \tfrac{1}{2}(J/I)^{\frac{1}{2}}. \end{aligned}\right\} \tag{3.27}$$

Hence the wave elevation may be expressed as

$$y-\eta = \epsilon\zeta = \frac{\epsilon J\phi_n}{K}\,\mathrm{sech}^2\left\{\frac{(\epsilon J)^{\frac{1}{2}}x}{2I^{\frac{1}{2}}}\right\}, \tag{3.28}$$

and the amplitude parameter ϵ may, in a specific physical application, be conveniently determined by the maximum of $|y-\eta|$. As explained just above, ϕ_n in this expression should be taken as a function of y rather than η for applications to internal waves ($n \geqslant 1$) when b is small. The fact that $K = O(b)$ for $n = 1$ now confirms the statement made in §1 that $l^{-2} = O(b\epsilon)$ for internal solitary waves.

It is particularly noteworthy that internal solitary waves are, like ones in homogeneous fluids, always supercritical; that is, the wave speed $|c|$ always exceeds the speed of infinitesimal long waves. The existence in theory of a variety of internal solitary waves corresponding to $n = 1, 2, 3, \ldots$, a sequence having no limit when $U = 0$, is probably not very significant in practical respects, since it would be difficult to produce the higher modes individually and they do not, of course, act independently of each other if superposed. The solitary wave corresponding to $n = 1$ is the most significant because, having a greater speed than any other internal-wave mode, it may emerge distinctly as the effects of a localized disturbance are propagated away from the origin. The author is currently undertaking some experiments on this aspect and hopes to report the findings in due course.

3.7. *Cnoidal waves*

Equation (3.22) has a periodic real solution when the cubic in f on the right-hand side has three distinct real roots. The supercritical and subcritical cases require separate discussion as follows. For clarity we assume that $K > 0$, but the physical conclusions are readily seen to be the same when $K < 0$.

Supercritical case

We have $J > 0$ and so the form of the cubic is as illustrated in figure 2(*a*). When $r-s = 0$, curve A is described; i.e. the cubic has double roots $f = 0$. Over the range where the cubic is positive between $f = 0$ and the higher root, equation (3.22) has the solitary-wave solution (3.27). For a periodic solution the curve must be lowered to become one of type B, and so it is necessary that $r-s > 0$. Thus, in physical terms, we see that periodic waves can arise from a supercritical flow only if some energy is lost, but at the same time not too much flow force is lost (e.g. by frictional drag at the fixed boundaries of the flow). The solution in this case represents an undular bore or hydraulic jump, and the physical conditions just mentioned are the precise counterpart of the conditions that Benjamin & Lighthill (1954) showed to apply to the undular bore in open-channel flows of homogeneous liquid.

If the three roots are denoted by f_1, f_2 and f_3 as shown in figure 2(*a*), with $f_1 > f_2 > 0$ and $f_3 < 0$ essentially, the solution of (3.22) is easily shown to be

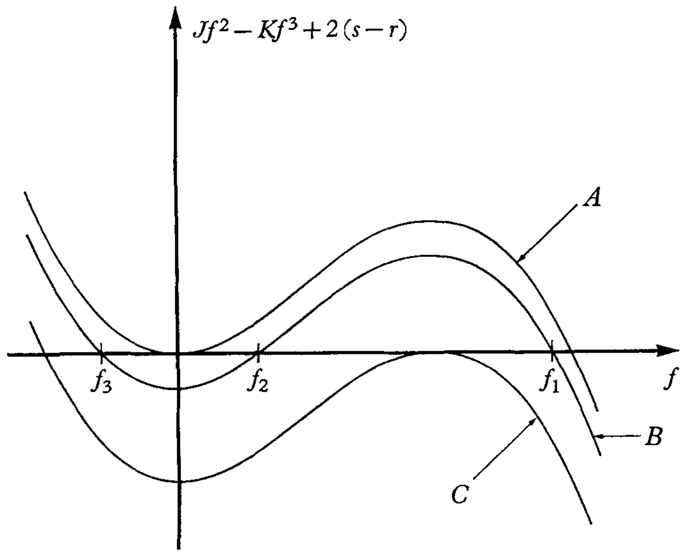
$$\left.\begin{aligned} f &= f_2 + (f_1 - f_2)\,\mathrm{cn}^2(mX; k), \\ m &= \left\{\frac{K(f_1-f_3)}{4I}\right\}^{\frac{1}{2}}, \quad k^2 = \frac{f_1-f_2}{f_1-f_3}. \end{aligned}\right\} \tag{3.29}$$

with

(cf. Lamb 1932, §253). Since the period of the elliptic function cn^2 is $2\mathrm{K}(k)$, the wavelength is $2m^{-1}\mathrm{K}(k)$. Note that the modulus $k \to 1$ as $-f_3 \to f_2 \to 0$ (curve A); in the limit the wavelength becomes infinite and, as expected, (3.29) reduces to the solitary-wave solution (3.27). If the cubic is lowered sufficiently (i.e. $r-s$ increased sufficiently) so that curve C is approached in figure 2(*a*), we have $f_2 \to f_1$ and so $k \to 0$; (3.29) then describes a sinusoidal wave of infinitesimal amplitude and wavelength π/m. This case may conveniently be left for discussion later.

Subcritical case

We now have $J < 0$ and so the form of the cubic is as shown in figure 2(*b*). In this case there is no non-trivial solution corresponding to curve A' for $s-r = 0$. Thus, no wave can arise from a subcritical parallel flow without change of energy or flow force, a conclusion already known for open-channel flows of homogeneous liquid (Benjamin & Lighthill 1954; Benjamin 1956). To produce waves in this case the curve must be *raised* to become of type B', and so we must have $s-r > 0$.



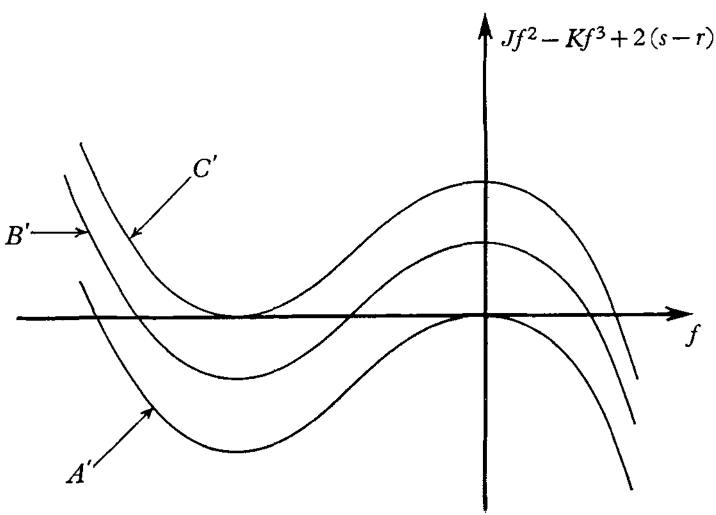


FIGURE 2. Forms of the cubic $Jf^2 - Kf^3 + 2(s-r)$ with $K > 0$: (*a*) supercritical case $J > 0$; (*b*) subcritical case $J < 0$.

This result represents the phenomenon of wave resistance: if a fixed obstacle spanning a subcritical flow generates a wave-train in its wake (lee waves), the drag on the obstacle equals the flow-force reduction $\epsilon^3 s$ in the receding stream. We also conclude that energy dissipation, measured by $\epsilon^3 r$, tends to diminish the wave amplitude. The cnoidal-wave solution (3.29) applies, with obvious modifications, to this case also.

When $s-r$ is made sufficiently large, the curves of type B' in figure 2 (*b*) approach curve C', for which the minimum point lies exactly on the axis. Curve C′ has the same essential features as curve A in figure 2 (*a*), except that the origin is shifted, and so the double root represents a supercritical parallel flow. As C' is approached from A', the waves described by the solution of (3.22) become progressively larger in amplitude and longer, tending finally to resemble a succession of solitary waves. The value of $\epsilon^3 s$ required to raise the cubic from A' to C' represents the largest wave resistance that can be manifested, in the absence of dissipation, by wave formation in the particular mode n on the given subcritical flow.

3.8. *Internal hydraulics*: *conjugate flows*

It must be remembered that the present theory is merely a first approximation to long waves of finite amplitude, and the flow state is assumed to be close to the critical in all examples (cf. equation (3.17)). Benjamin (1962*b*, § 3.2) has shown in general how a number N may be defined such that $N > 1$ represents supercritical states and $N < 1$ subcritical ones. [One considers travelling-wave disturbances superposed on a given steady flow, and lets c_+ and c_- denote the absolute velocities, measured positively in the direction of flow, at which waves of extreme length propagate respectively with and against the flow. Then this characteristic number may be defined by $N = (c_+ + c_-)/(c_+ - c_-)$.] Accordingly, the present condition is expressible as $N = 1 + O(b\epsilon)$, where $b\epsilon$ is small by our basic hypothesis. However, several highly plausible physical conclusions regarding flow states with N further from 1 may now be drawn by analogy with the results of the classical theory of open-channel flows. At the same time we may suitably introduce into the discussion the principle of 'conjugate-flow pairs' which was established in a general way by Benjamin (1962*b*) with special regard to swirling flows, but whose application to heterogeneous-fluid flows is equally valid and was pointed out as such in the previous paper. The following physical arguments are intended primarily to refer to the first internal-wave mode corresponding to $n = 1$. Similar interpretations might be conjectured for the higher modes, though with less certainty for the reasons already noted in connexion with the solitary-wave solution.

First, we may suppose that as N is progressively raised above 1 the internal solitary wave will approach an extreme form, corresponding to the sharp-crested form that is a well-known result for the classical solitary wave of maximum height (which occurs at $N = 1{\cdot}25$ approximately). An attempt to produce a solitary wave with N greater than the limiting value, say N_c, will lead to 'breaking' of the wave, presumably some kind of mixing process inside the fluid with accompanying dissipation of energy.† We may further conclude that a bore

† So far very little seems to be known experimentally about the breaking of progressive internal waves, though recent experiments by Mr S. A. Thorpe at Cambridge have greatly illuminated the process of breaking for standing oscillations in stratified liquids. The nature of breaking in the first case must obviously be something quite different from the formation of white-caps by ordinary water waves.

with N significantly greater than N_c will break at the front, since a *small* loss of energy at the front now cannot produce periodic waves in the way previously explained. This is a well-known feature of bores in ordinary open-channel flows, which develop turbulent fronts when $N > 1{\cdot}25$ and become progressively more turbulent as N is raised.

Waves of extreme form can also be produced from a subcritical flow, when N is sufficiently smaller than 1 and the flow-force reduction $\epsilon^3 s$ is large enough. That is, a change like $A' \rightarrow C'$ in figure 2 (*b*) can no longer be brought about by the smooth formation of waves; in the sequence of possible cases given by increasing $\epsilon^3 s$ step by step, breaking intervenes before a state neighbouring on the critical one C' is reached. Thus, when N for the oncoming flow is below but not close to 1, lee waves formed behind an obstacle can be swelled to breaking if the size of the obstacle, and hence the wave drag upon it, is made large enough (cf. Benjamin 1956, pp. 231, 245)†.

We have that the change $A' \rightarrow C'$, or again $C \rightarrow A$, implies a reduction in flow force. But the minimum points of A and C' represent supercritical uniform flows, while the maximum points of C and A' represent subcritical ones. Hence we conclude that any supercritical flow with N near 1 possesses an energy-conserving 'conjugate' state for which the flow force is larger. This is in fact a general result unrestricted as to the value of N; it was proved by Benjamin (1962*b*) using a variational argument and is well known in the simple instance of ordinary open-channel flows. Benjamin also proved in general that, as the value of N for a supercritical flow is raised by varying some physical parameter, the conjugate subcritical value of N is lowered and the flow-force excess of the conjugate state is increased. These properties may readily be confirmed from the present results. Our extension of the conjugate-flow principle to include the effects of small energy losses explicitly is special to the case of near-critical conditions, however; losses that are not small (i.e. when H changes by large fractions) cannot properly be analysed without additional physical hypotheses.

The change from the subcritical to the supercritical member of a conjugate

† In this, as in all other instances where a steady disturbance is supposed to be generated in a *subcritical* flow, the following important point of interpretation needs to be borne in mind. When an obstacle is introduced into a pre-existing subcritical flow, its effect will generally propagate far upstream so that, after a steady state is reached, the flow prevailing in front of the obstacle will not be the same as the original one. This phenomenon is commonly called 'blocking' and it reflects the general principle, well known in open-channel hydraulics, that subcritical states of flow depend essentially on the conditions imposed downstream and so cannot be determined arbitrarily at their sources. Accordingly, in a steady-state model for lee waves, the parallel flow specified in front of the obstacle must itself, as well as the waves downstream, be regarded as a development from some other, more basic flow that would occur in the absence of the obstacle. The upstream influence of a small obstruction is often quite small in comparison with the wave effects produced downstream, however, and so it may be reasonable to investigate the downstream effects of a varying obstruction on the basis of the same model for the flow upstream. Moreover, this somewhat awkward point of physical interpretation is effectively circumvented by any theory that, like the present one, allows the primary density and velocity distributions to take arbitrary functional forms; the complete gamut of possible flow states upstream and downstream of an obstacle is thus covered in principle, independently of practical questions regarding the realizability of any particular flow upstream.

flow pair, as illustrated by $A' \to C'$ in figure 2(*b*), represents a situation that might well be realized in practice if the subcritical N is not too large. The supercritical flow may form downstream from an obstacle large enough to exert the large force $S_0 - S$, provided the conditions far downstream are favourable; specifically, they must not impede the flow to the extent of precipitating an internal hydraulic jump immediately behind the obstacle (Benjamin 1965, p. 521). The situation in view is analogous to the one in open-channel flow where a transition from subcritical to supercritical states is brought about by a sluice gate or other obstacle spanning the stream.

The change corresponding to $A \to C$ in figure 2(*a*) is the counterpart of the classical model for a dissipative bore (Lamb 1932, §280), in which a transition from a supercritical uniform flow to a subcritical one is supposed to occur for a constant flow force. Let us also assume $s = 0$. The value of $r\,(>0)$ required to lower the cubic from A to C represents the maximum dissipation then possible at the front of the bore; and if r is less than this value the bore must be undular in form. The wave amplitude is obviously largest when r is very small (provided of course that $1 < N < N_c$ and therefore waves nearly the same as solitary waves are possible), and the amplitude decreases to zero as the dissipation approaches the maximum and a uniform subcritical flow is produced downstream.

3.9. *The special properties of linear systems*

It was explained at the end of §2 that, if the basic equation (2.13), or (2.8), for the stream-functions is linear, the expansion of $S - S_0$ terminates at second order in ϵ when the upper boundary is fixed. Hence in any such case we shall find $K = 0$, so that there is no solitary-wave solution. This conclusion is obvious from the flow-force integral expressed in the form (2.12), but it is not immediately evident from the alternative form (3.3) which is the basis of the present analytical development. However, since the two forms are equivalent and since the existence of solitary waves obviously cannot depend on the choice between $\psi(x, y)$ or $y(x, \eta)$ as dependent variable, we are bound to find for every case in question that $K = 0$ according to the definition (3.25). The reason why the conclusion holds only for a fixed upper boundary is clear: the boundary condition (3.7) at a free surface is non-linear and can provide the non-linear ingredients essential to solitary waves, just as in the classical theory where the equation for the stream-function is of course linear (Laplace's equation). Nevertheless, it is remarkable that only the linearized form (3.8) of the free-boundary condition enters the present analysis explicitly, though the non-linear effects of a free surface are implicit in the condition of constant flow force as used in the final approximation.

The non-existence of a solitary-wave solution implies there is no neighbouring cnoidal-wave solution derivable for a supercritical flow subject to energy loss. Thus no steady bore or hydraulic jump is possible in a linear system. This conclusion was demonstrated in a quite different way by Benjamin (1962*b*) with regard to vortex flows. It must be recognized, however, that in general both the density and velocity distributions, ρ and W, must be chosen in special ways to give a linear system (see Long 1958; Yih 1960*b*). Accordingly, only the possibility of a *stationary* hydraulic jump, i.e. in a situation for which the velocity of the flow upstream is a basic specification of the physical model, is ruled out by the artifice of making the system linear. For travelling bores, on the other hand, only ρ and U can be regarded as basic specifications; $W = c+U$ is not, because c is a wave property. A choice of W to make the respective steady-flow model linear is not necessarily allowable, therefore, and presumably travelling solitary waves and bores are possible for any given ρ and U, since they are free to travel at speeds that make the relative flow a non-linear system.

Only one relevant example comes to mind in which finite-amplitude waves of permanent form are known *a priori* to satisfy a linear equation. This is the case of axisymmetric disturbances in a fluid with constant density and with uniform angular velocity together with zero axial velocity in the undisturbed state. In keeping with the present conclusions, Benjamin & Barnard (1964) have shown both theoretically and experimentally that a continuous disturbance propagated into the steady-state region cannot develop into a steady bore, but is instead headed by a continually dispersing wave-front.

The foregoing remarks only concern applications to supercritical flows, and on the basis of present ideas there is no objection to the use of linear models for lee waves, which necessarily arise from subcritical flows. When $K = 0$, the cubic in f represented in figure 2 reduces to the quadratic $Jf^2+2(s-r)$, which is curved downwards in the subcritical case ($J < 0$). Thus periodic solutions are still possible when the wave-resistance parameter $s > 0$. The solution of (3.22) is then sinusoidal, as indeed must be expected since finite-amplitude waves satisfy the same equations as infinitesimal ones when the overall system is linear.

An important qualitative distinction remains, however, between subcritical flows in linear and non-linear systems. It is that for the former no conjugate supercritical flow exists which could be produced from the subcritical state by a large reduction in flow force (i.e. large s). This other peculiarity of linear systems was also pointed out by Benjamin (1965), and it is demonstrated now by the fact that the parabola replacing the cubic curves in figure 2(*b*) has no minimum point which, by a sufficient increase of s, could be raised to touch the axis.

Allied to the use of linear systems to provide tractable models of stratified-flow phenomena, use has often been made of certain approximations, notably the so-called Boussinesq approximation (e.g. see Long 1965), which may be justified on the assumption that the density gradient is sufficiently small. The present results incidentally give some helpful insight into the significance of such approximations, though we shall not digress to discuss the matter here. It is taken up in the Appendix to this paper.

4. Examples

The chief advantage of the preceding analysis is that it gives a general account of internal waves of finite amplitude and permanent form while avoiding the need to focus attention on any particular physical model. The leading results comprising equation (3.22) and its solutions are in forms that may readily be applied to any example, however, and they could even be used without undue difficulty for applications where the primary density and velocity distributions are determined experimentally. It would be required only to solve the ordinary differential equation (3.14), which might be done readily enough by machine computation, and then to evaluate the integrals (3.23)–(3.25).

The properties of the solitary wave will now be worked out in three examples to illustrate the application of the theory and to provide some comparisons with results found previously in other ways. It seems justified to claim that the present general formulae give the wave speed and other properties with greater facility than any other known method.

Example 1. *The classical solitary wave*

This comprises a degenerate case of the present results, given by putting $\rho = \rho_0$ (const.), $Q = \rho_0 c^2$. The upper boundary must, of course, be free; there is obviously no solution when the boundary is fixed.

In this case equation (3.14) for the eigenfunction ϕ_0 reduces to

$$\frac{d^2\phi_0}{d\eta^2} = 0,$$

and the solution satisfying the lower boundary condition is

$$\phi_0 = \eta.$$

The upper boundary condition (i.e. the last equation of (3.13) with $c = c_0$) now gives

$$c_0^2 = gh_0, \tag{4.1}$$

which is the familiar classical result for the speed of infinitesimal long waves.

We have

$$\epsilon\gamma_0 = Q - Q_0 = \rho_0(c^2 - c_0^2),$$

and, for the coefficients defined by (3.23)–(3.25),

$$I = \int_0^{h_0} \rho_0 c^2 \eta^2 d\eta = \tfrac{1}{3}\rho_0 c^2 h_0^3,$$

$$\epsilon J = \rho_0(c^2 - c_0^2)\, h_0,$$

$$K = \rho_0 c_0^2 h_0.$$

Hence substitution into (3.28) gives

$$y - \eta = \eta\left(\frac{c^2}{c_0^2} - 1\right)\operatorname{sech}^2\left\{\frac{(3c^2 - 3c_0^2)^{\frac{1}{2}} x}{2ch_0}\right\}. \tag{4.2}$$

For the elevation $y(h_0) - h_0$ of the free surface, denoting its maximum value by Δ, we get from (4.1) and (4.2)

$$y(h_0) - h_0 = \Delta \operatorname{sech}^2\left\{\left(\frac{3\Delta}{h_0 + \Delta}\right)^{\frac{1}{2}} \frac{x}{h_0}\right\}, \tag{4.3}$$

and

$$c^2 = g(h_0 + \Delta). \tag{4.4}$$

These are precisely the results found originally by Rayleigh (see Lamb 1932, §252) and confirmed by many other writers since his time.

Example 2. *Two-layer system*

The physical model is illustrated in figure 3. Take $\rho = \rho_1$ (const.) in $0 \leqslant \eta < h_1$, and $\rho = \rho_2$ (const. $< \rho_1$) in $h_1 < \eta \leqslant h_0 = h_1 + h_2$. Also take the upper boundary $\eta = h_0$ to be a rigid plane, and assume $U = 0$ so that $Q = \rho c^2$. Put $R = h_1/h_0$ and $\sigma = \rho_2/\rho_1$. The problem of the solitary wave in this system has been treated by Keulegan (1953) and, in more detail, by Long (1956).

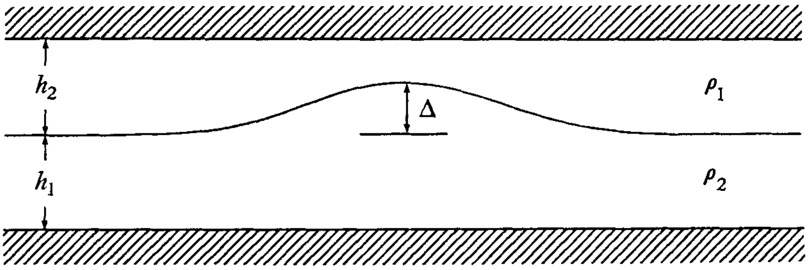


FIGURE 3. Illustration of two-layer system showing maximum vertical displacement Δ of interface by solitary wave.

There is only one possible long-wave mode, and so the required solution of (3.14) may be written ϕ without need of a suffix. The respective critical wave speed is denoted by $\bar{c}$. Equation (3.14) becomes

$$\frac{d^2\phi}{d\eta^2} = 0$$

in each layer separately, and the solution satisfying the boundary conditions $\phi(0) = \phi(h_0) = 0$ and being continuous at the interface $\eta = h_1$ (as required for continuity of the vertical component of velocity) is

$$\left.\begin{aligned} \phi &= \eta && \text{in} \quad 0 \leqslant \eta < h_1, \\ &= \frac{R}{1-R}(h_0 - \eta) && \text{in} \quad h_1 < \eta \leqslant h_0. \end{aligned}\right\} \tag{4.5}$$

Note that, since ϕ is largest at the interface, the maximum vertical displacement at a given x occurs there.

Applying the limit process

$$\lim_{\delta \to 0} \int_{h_1 - \delta}^{h_1 + \delta} \{ \ \} \, d\eta$$

to equation (3.14), we find a second condition at the interface to be

$$\bar{c}^2 \langle \rho \, d\phi/d\eta \rangle = g \langle \rho\phi \rangle, \tag{4.6}$$

where $\langle \ \rangle$ denotes the difference in the enclosed quantity across the interface. This in fact represents the requirement that the pressure is continuous, and it is seen to imply that the horizontal component of velocity is discontinuous. Substitution of (4.5) into (4.6) leads to

$$\bar{c}^2 = \frac{(1-R)(1-\sigma)}{1-R+\sigma R} g h_1, \tag{4.7}$$

which agrees with the formula given by Lamb (1932, p. 371) for the speed of infinitesimal waves.

Equations (3.23)–(3.25) now give

$$I = \tfrac{1}{3}\rho_1 c^2 h_1^3 \left(1-\sigma+\frac{\sigma}{R}\right),$$

$$\epsilon J = \rho_1 (c^2-\bar{c}^2) h_1 \left(1+\frac{\sigma R}{1-R}\right),$$

$$K = \rho_1 \bar{c}^2 h_1 \left\{1-\sigma\left(\frac{R}{1-R}\right)^2\right\}.$$

Substituting these expressions into the solitary-wave solution (3.28), we obtain for the displacement of the interface

$$y(h_1)-h_1 = \Delta\,\mathrm{sech}^2\,\omega x, \tag{4.8}$$

with

$$\frac{c^2}{\bar{c}^2} = 1+\frac{(1-R)^2-\sigma R^2}{(1-R)(1-R+\sigma R)}\frac{\Delta}{h_1}, \tag{4.9}$$

$$\omega^2 = \frac{3\Delta R\{(1-R)^2-\sigma R^2\}}{4h_1^3(1-R)^2(R+\sigma-\sigma R)}\frac{\bar{c}^2}{c^2}. \tag{4.10}$$

These results agree with those found by Keulegan (1953) and Long (1956, cf. equations (20) and (28) with $F_1 = F_2$). They reduce to the results (4.3), (4.4) in the previous example if the density of the upper liquid is made vanishingly small, i.e. $\sigma \to 0$.

Note that the wave is one of elevation ($\Delta > 0$) or depression ($\Delta < 0$) accordingly as

$$(1-R)^2-\sigma R^2 \gtrless 0 \tag{4.11}$$

(cf. Long, equations (22)). These conditions are found to be equivalent to

$$R \lessgtr \tfrac{1}{2}+\tfrac{1}{8}(1-\sigma)+\tfrac{1}{16}(1-\sigma)^2+O(1-\sigma)^3.$$

Thus, if the fractional density difference $1-\sigma = (\rho_1-\rho_2)/\rho_1$ is very small, the wave is one of elevation when the upper layer is the deeper of the two, and is one of depression when the lower layer is the deeper.

Example 3. *Fluid with exponential density gradient*

The physical model is illustrated in figure 4. It has previously been investigated by Peters & Stoker (1960) in the case of a free upper boundary, and by Long (1965) and Benney (1966) in the case of a fixed upper boundary. We take

$$\rho = \rho_0 e^{-\beta\eta} (\beta > 0) \quad \text{and} \quad U = 0,$$

so that $Q = \rho_0 c^2 e^{-\beta\eta}$.

Equation (3.14) now becomes

$$\frac{d^2\phi_n}{d\eta^2}-\beta\frac{d\phi_n}{d\eta}+\frac{g\beta}{c_n^2}\phi_n = 0, \tag{4.12}$$

and the solution satisfying the lower boundary condition $\phi_n(0) = 0$ is

$$\phi_n = e^{\frac{1}{2}\beta\eta}\sin\lambda_n\eta, \tag{4.13}$$

with

$$\lambda_n^2 = \frac{g\beta}{c_n^2}-\frac{\beta^2}{4}. \tag{4.14}$$

From this point on, the cases of a fixed and of a free upper boundary require separate treatment.

Fixed upper boundary

The boundary condition $\phi_n(h_0) = 0$ shows that

$$\lambda_n h_0 = n\pi \quad (n = 1, 2, 3, \ldots). \tag{4.15}$$

Hence (4.14) gives

$$c_n^2 = \frac{g\beta h_0^2}{n^2\pi^2 + \frac{1}{4}\beta^2 h_0^2}. \tag{4.16}$$

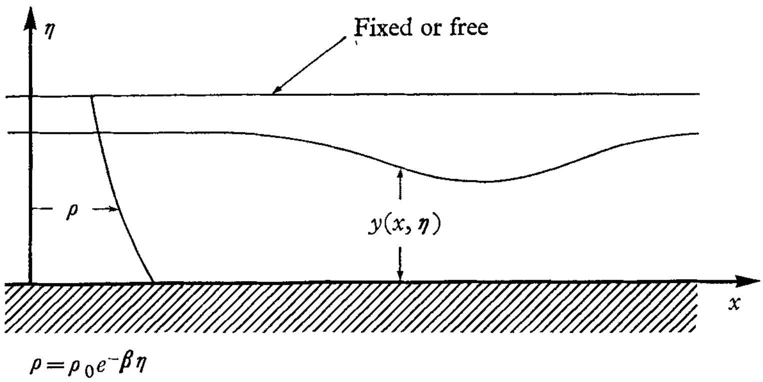


FIGURE 4. Illustration of flow system in which the density ρ of the fluid decreases exponentially with height.

The coefficients defined by (3.23)–(3.25) are found to be

$$I = \rho_0 c^2 \int_0^{h_0} \sin^2 \lambda_n \eta \, d\eta = \tfrac{1}{2}\rho_0 c^2 h_0, \tag{4.17}$$

$$\epsilon J = \rho_0(c^2 - c_n^2) \int_0^{h_0} (\lambda_n \cos \lambda_n \eta + \tfrac{1}{2}\beta \sin \lambda_n \eta)^2 \, d\eta$$

$$= \tfrac{1}{2}\rho_0(c^2 - c_n^2)\, h_0^{-1}(n^2\pi^2 + \tfrac{1}{4}\beta^2 h_0^2), \tag{4.18}$$

$$K = \rho_0 c_n^2 \int_0^{h_0} e^{\frac{1}{2}\beta\eta}(\lambda_n \cos \lambda_n \eta + \tfrac{1}{2}\beta \sin \lambda_n \eta)^3 \, d\eta$$

$$= \frac{\rho_0 c_n^2 \beta n^3 \pi^3 \{1 - (-1)^n \exp(\frac{1}{2}\beta h_0)\}}{h_0(9n^2\pi^2 + \frac{1}{4}\beta^2 h_0^2)}. \tag{4.19}$$

Substituting (4.13) and (4.17)–(4.19) into (3.28), we are led finally to

$$y - \eta = (-1)^{n+1} \Delta e^{\frac{1}{2}\beta\eta} \sin\left(\frac{n\pi y}{h_0}\right) \operatorname{sech}^2 \omega x, \tag{4.20}$$

with $\Delta > 0$ and

$$\frac{c^2}{c_n^2} = 1 + \frac{2n^3\pi^3\beta\Delta\{\exp(\frac{1}{2}\beta h_0) + (-1)^{n+1}\}}{(n^2\pi^2 + \frac{1}{4}\beta^2 h_0^2)(9n^2\pi^2 + \frac{1}{4}\beta^2 h_0^2)}, \tag{4.21}$$

$$\omega^2 h_0^2 = \frac{n^3\pi^3 c_n^2 \beta\Delta\{\exp(\frac{1}{2}\beta h_0) + (-1)^{n+1}\}}{2c^2(9n^2\pi^2 + \frac{1}{4}\beta^2 h_0^2)}. \tag{4.22}$$

These results agree with those found by Long (1965) and Benney (1966). Note that the solitary wave for $n = 1$ is a wave of elevation in this case.

On the assumption that $\beta h_0 \ll 1$, which covers most practical applications, (4.16) gives as a good first approximation

$$c_n^2 = \frac{g\beta h_0^2}{n^2\pi^2}, \tag{4.23}$$

while (4.21) and (4.22) give approximately

$$\left.\begin{aligned} \frac{c^2}{c_n^2} &= 1+\frac{4\beta\Delta}{9n\pi}, \quad n \text{ odd}, \\ &= 1+\frac{\beta^2 h_0 \Delta}{9n\pi}, \quad n \text{ even}, \end{aligned}\right\} \tag{4.24}$$

$$\left.\begin{aligned} \omega^2 h_0^2 &= \frac{n\pi\beta\Delta}{9}, \quad n \text{ odd}, \\ &= \frac{n\pi\beta^2 h_0 \Delta}{36}, \quad n \text{ even}. \end{aligned}\right\} \tag{4.25}$$

It is interesting that for n even the values of $c^2 - c_n^2$ and ω^2 are greatly less, by a factor $O(\beta h_0)$, than the values for n odd. This curious property was noted by Long (1965). The results for n even seem unlikely to have any practical significance, however. For a very small βh_0, as is usual, a solitary wave in this category is so extremely long, and its speed so little different from c_n, that its detection as a true wave of permanent form is probably impossible.

Free upper boundary

The boundary condition on ϕ_n is now

$$c_n^2 \frac{d\phi_n}{d\eta} = g\phi_n \quad \text{at} \quad \eta = h_0, \tag{4.26}$$

which, when (4.13) is substituted, gives

$$\lambda_n \cot \lambda_n h_0 + \tfrac{1}{2}\beta = g/c_n^2. \tag{4.27}$$

We shall henceforth exclude the case $n = 0$, for which (4.27) is satisfied by (4.1) very nearly (with λ_0 very small according to (4.14)).

On the assumption that βh_0 is small, (4.14) shows the right-hand side of (4.27) to be large in comparison with $\lambda_n (n > 0)$. Hence we deduce that $\lambda_n h_0 - n\pi$ is a small positive number, and to a first approximation we find that

$$\lambda_n h_0 = n\pi + \frac{\beta h_0}{n\pi}, \tag{4.28}$$

$$c_n^2 = \frac{g\beta}{\lambda_n^2}. \tag{4.29}$$

Thus the infinitesimal-wave speed is very nearly the same as when the upper boundary is fixed, and the only qualitative difference in the solution is that ϕ_n changes sign just below the surface, so that in a thin layer at the top the vertical displacement of the fluid is in a direction opposite to that below.

Since the solitary-wave results become unduly complicated otherwise, only the first approximation for small βh_0 will be given. The expressions (4.17) and (4.18) therefore suffice for I and ϵJ. In the evaluation of K from the integral preceding (4.19), the only difference is that, after transformation to the dimensionless variable of integration $\lambda_n \eta$, the upper limit is given by (4.28) instead of (4.15) as before. It is found in this case that

$$\left.\begin{aligned} K &= -\tfrac{7}{9}\rho_0 c_n^2 \lambda_n \beta, \quad n \text{ odd}, \\ &= \rho_0 c_n^2 \lambda_n \beta, \quad n \text{ even}. \end{aligned}\right\} \tag{4.30}$$

Hence, in place of (4.20), (4.24) and (4.25), the results are

$$y-\eta = (-1)^n \Delta e^{\frac{1}{2}\beta\eta} \sin\left\{\left(\frac{n\pi}{h_0}+\frac{\beta}{n\pi}\right) y\right\} \operatorname{sech}^2 \omega x, \tag{4.31}$$

$$\left.\begin{aligned} \frac{c^2}{c_n^2} &= 1+\frac{14\beta\Delta}{9n\pi} \quad n \text{ odd}, \\ &= 1+\frac{2\beta\Delta}{n\pi}, \quad n \text{ even}, \end{aligned}\right\} \tag{4.32}$$

$$\left.\begin{aligned} \omega^2 h_0^2 &= \frac{7n\pi\beta\Delta}{18}, \quad n \text{ odd}, \\ &= \frac{n\pi\beta\Delta}{2}, \quad n \text{ even}. \end{aligned}\right\} \tag{4.33}$$

Thus the solitary wave for $n = 1$ is a wave of depression in this case, except in a thin layer just below the surface where the displacements are in the upward direction.

This major distinction between the properties of solitary waves according to whether the upper boundary is fixed or free appears very remarkable, particularly since at first sight one might reasonably suppose the effect of the difference in boundary conditions to become insignificant for internal waves when β is sufficiently small. This is a convincing demonstration of the principle, recently discussed by Long (1965), that small effects without obvious importance and such as might be neglected in certain approximate theories may in fact be crucial with regard to finite-amplitude waves (see Appendix).

The present formulae (4.32) and (4.33) agree with the results found by Peters & Stoker (1960) for $n = 1$.

5. Conclusion

Perhaps the most obvious limitation of the main analysis in § 3 is the absence of any proof of convergence for the method of approximate solution. It is, in effect, assumed that waves of permanent form exist and then their properties are worked out approximately, but of course this approach does not establish their reality with complete certainty. The author believes, however, that this aspect is already adequately covered by the work of Ter-Krikorov (1963), which justifies complete confidence in the propriety of first-order approximations to permanent internal waves. Without this analytical background there might be serious cause for doubt because, unlike the situation as regards the ordinary solitary wave, no conclusive experimental evidence is available as to the permanence of internal solitary waves.

It seems likely that an existence proof on simpler lines than Ter-Krikorov's argument could be constructed on the basis of the present method of approach, proceeding from the flow-force integral (3.3). An encouraging fact in this regard is that, since the integral has fixed limits and the only x-derivative present is cleared in the first approximation, there would be little difficulty in finding definite bounds on the error in any approximate evaluation of the integral. Hence proof of the existence of a solution by a method of the successive-approximations type seems feasible. This situation contrasts markedly with the very awkward one presented by orthodox methods of approach to solitary-wave theory (see Lamb 1932, §252), where progressively higher derivatives occur explicitly in each successive stage of approximation. Moreover, the physical meaning of the generalized solitary-wave theory, including cnoidal-wave solutions, is completely clarified by the present approach. The physical circumstances (e.g. the possible modes of generation) of periodic waves of permanent form remain highly obscure in the absence of an explicit definition of them in terms of flow force (cf. Lamb, §253).

The present theory could be applied more or less intact to axisymmetric swirling flows, which are well known to be mathematically equivalent to two-dimensional stratified flows. Conversely, theory developed in this other context is equally meaningful here. For instance, the author's theory of the vortex breakdown phenomenon (Benjamin 1962*b*, 1965) can be used to prove the possibility of undular internal bores with complete generality as regards the primary density and velocity distributions, and it provides a basic framework (akin, one might say, to the basic theory of the hydraulic jump in open-channel flow) into which the present results fit quite naturally (in much the same way as Benjamin & Lighthill's (1954) presentation of classical solitary- and cnoidal-wave theory was related to the hydraulic jump). The possibility of large periodic disturbances arising from a supercritical swirling flow, after some slight loss of energy occurs, was argued as an essential of the vortex breakdown theory, and though the context is different the present results virtually demonstrate the solutions that describe such disturbances.

Appendix. Note on the Boussinesq approximation

This approximation is commonly used in studies of internal waves, being justifiable in some applications where the fractional density variations are very small. The effect of the density variations on the inertial terms in the equations of motion is neglected, density being assigned its (constant) mean value where it appears as a factor in these terms; but density is still represented precisely as a function of height where it appears multiplied by g in the equation of motion for the vertical direction.

The significance of the Boussinesq approximation has recently been discussed by Long (1965). He showed that its use generally demands caution, particularly when waves of finite amplitude are in question, and he pointed out in a specific example that the internal solitary wave depends crucially upon small effects neglected in this approximation. The matter may be illustrated very readily from the preceding analytical results, and so a few points adding to Long's discussion will now be noted.

In the present notation, the Boussinesq approximation amounts to putting $\rho = \text{const.}$ in the function Q. If for simplicity of illustration we limit attention to the case $U = 0$ (i.e. waves propagating towards liquid at rest), we then have $Q = \rho c^2 = \text{const.}$ as the approximation to be examined. We also limit attention to the case of a fixed upper boundary.

First consider equation (3.12) for infinitesimal sinusoidal waves. On the substitution of $\phi = Q^{-\frac{1}{2}}\Phi$, the equation becomes

$$\Phi_{\eta\eta} + \left\{-\alpha^2 - \frac{g\rho_\eta}{Q} - \frac{1}{2}\frac{Q_{\eta\eta}}{Q} + \frac{1}{4}\frac{Q_\eta^2}{Q^2}\right\}\Phi = 0. \tag{A1}$$

One now sees clearly that, if the fractional variation of ρ is indeed small on the scale of the overall depth, then the approximation $Q = \text{const.}$ has an insignificant effect on the solution Φ, and hence ϕ. Because of the boundary conditions $\Phi(0) = \Phi(h_0) = 0$, the coefficient $\{\ \}$ in (A1) is $O(n\pi/h_0)^2$, being raised to this order of magnitude by the smallness of $Q/g = \rho c^2/g$ as compared with ρ_η; and hence the final two terms in the coefficient are justifiably negligible. Thus, for infinitesimal waves at least, the Boussinesq approximation appears self-consistent and to have no qualitative effect on the solutions obtained.

On checking its consequences in our non-linear theory, however, we see the possibility of serious error. Whereas the coefficients I and J of the general equation (3.22) for long waves are found to be little affected by the approximation, the coefficient K appears to be altered radically by it. Integrating the right-hand side of (3.25) by parts and using (3.14), we obtain

$$\begin{aligned} K &= -\int_0^{h_0} \phi_n \frac{d\phi_n}{d\eta}\left(2Q_n\frac{d^2\phi_n}{d\eta^2} + \frac{dQ_n}{d\eta}\frac{d\phi_n}{d\eta}\right)d\eta \\ &= -\int_0^{h_0} \phi_n \frac{d\phi_n}{d\eta}\left(2g\frac{d\rho}{d\eta}\phi_n - \frac{dQ_n}{d\eta}\frac{d\phi_n}{d\eta}\right)d\eta \\ &= \int_0^{h_0}\left\{\tfrac{2}{3}g\frac{d^2\rho}{d\eta^2}\phi_n^3 + \frac{dQ_n}{d\eta}\phi_n\left(\frac{d\phi_n}{d\eta}\right)^2\right\}d\eta. \end{aligned} \tag{A2}$$

The two terms on the right-hand side of (A2) are generally of comparable size, but the Boussinesq approximation removes the second of them. Moreover, an additional approximation, often introduced in company with the Boussinesq approximation, is to take ρ to be a linear function of η, in which case the first term of (A2) disappears also. With $K = 0$ there is no solitary-wave solution.

When the density variation is small, the coefficient K is necessarily small, but an accurate evaluation of it is still essential to a reliable description of any solitary wave possible in the system.

REFERENCES

BENJAMIN, T. B. 1956 On the flow in channels when rigid obstacles are placed in the stream. *J. Fluid Mech.* **1**, 227.

BENJAMIN, T. B. 1962*a* The solitary wave on a stream with an arbitrary distribution of vorticity. *J. Fluid Mech.* **12**, 97.

BENJAMIN, T. B. 1962*b* Theory of the vortex breakdown phenomenon. *J. Fluid Mech.* **14**, 593.

BENJAMIN, T. B. 1965 Significance of the vortex breakdown phenomenon. *Trans. Amer. Soc. Mech. Engrs, J. Basic Eng.* **87**, 518.

BENJAMIN, T. B. & BARNARD, B. J. S. 1964 A study of the motion of a cavity in a rotating liquid. *J. Fluid Mech.* **19**, 193.

BENJAMIN, T. B. & LIGHTHILL, M. J. 1954 On cnoidal waves and bores. *Proc. Roy. Soc.* A, **224**, 448.

BENNEY, D. J. 1966 Long non-linear waves in fluid flows. To appear in *J. Math. Phys.*

DUBREIL-JACOTIN, M. L. 1937 Sur les théorèmes d'existence relatifs aux ondes permanentes périodiques à deux dimensions dans les liquides hétérogènes. *J. Math. Pures Appl.* (9), **19**, 43.

FRIEDRICHS, K. O. & HYERS, D. H. 1954 The existence of solitary waves. *Comm. Pure Appl. Math.* **7**, 517.

HOWARD, L. N. 1961 Note on a paper by John W. Miles. *J. Fluid Mech.* **10**, 509.

INCE, E. L. 1926 *Ordinary Differential Equations.* London: Longmans. (Dover edition 1956.)

JEFFREYS, H. & JEFFREYS, B. S. 1946 *Methods of Mathematical Physics.* Cambridge University Press.

KEULEGAN, G. H. 1953 Characteristics of internal solitary waves. *J. Res., Nat. Bureau of Standards,* **51**, 133.

KORTEWEG, D. J. & DE VRIES, G. 1895 On the change of form of long waves advancing in a rectangular canal and a new type of long stationary waves. *Phil. Mag.* (5), **39**, 422.

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

LITTMAN, W. 1957 On the existence of periodic waves near critical speed. *Comm. Pure Appl. Math.* **10**, 241.

LONG, R. R. 1953 Some aspects of the flow of stratified fluids. Part I. A theoretical investigation. *Tellus,* **5**, 42.

LONG, R. R. 1956 Solitary waves in one- and two-fluid systems. *Tellus,* **8**, 460.

LONG, R. R. 1958 Tractable models of steady-state stratified flow with shear. *Quart. J. Roy. Meteor. Soc.* **84**, 159.

LONG, R. R. 1964 The initial-value problem for long waves of finite amplitude. *J. Fluid Mech.* **20**, 161.

LONG, R. R. 1965 On the Boussinesq approximation and its role in the theory of internal waves. *Tellus,* **17**, 46.

MILES, J. W. 1961 On the stability of heterogeneous shear flows. *J. Fluid Mech.* **10**, 496.

MILES, J. W. 1963 On the stability of heterogeneous shear flows. Part 2. *J. Fluid Mech.* **16**, 209.

PETERS, A. S. & STOKER, J. J. 1960 Solitary waves in liquids having non-constant density. *Comm. Pure Appl. Math.* **13**, 115.

SHEN, M. C. 1964 Solitary waves in compressible media. *New York Univ. Inst. Math. Sci., Rep.* IMM-NYU 325.

SHEN, M. C. 1965 Solitary waves in running gases. *New York Univ. Inst. Math. Sci., Rep.* IMM-NYU 341.

TER-KRIKOROV, A. M. 1963 Théorie exacte des ondes longues stationnaires dans un liquide hétérogène. *J. Mécanique,* **2**, 351.

URSELL, F. 1953 The long-wave paradox in the theory of gravity waves. *Proc. Camb. Phil. Soc.* **49**, 685.

YANOWITCH, M. 1962 Gravity waves in a heterogeneous incompressible fluid. *Comm. Pure Appl. Math.* **15**, 45.

YIH, C.-S. 1960*a* Gravity waves in a stratified fluid. *J. Fluid Mech.* **8**, 481.

YIH, C.-S. 1960*b* Exact solutions for steady two-dimensional flow of a stratified fluid. *J. Fluid Mech.* **9**, 161.