J. Fluid Mech. (1978), *vol.* 85, *part* 1, *pp.* 7–31
Printed in Great Britain

On the shape and breaking of finite amplitude internal gravity waves in a shear flow

By S. A. THORPE
Institute of Oceanographic Sciences, Wormley, Godalming, Surrey, England

(Received 17 May 1977)

This paper is concerned with two important aspects of nonlinear internal gravity waves in a stably stratified inviscid plane shear flow, their shape and their breaking, particularly in conditions which are frequently encountered in geophysical applications when the vertical gradients of the horizontal current and the density are concentrated in a fairly narrow depth interval (e.g. the thermocline in the ocean). The present theoretical and experimental study of the wave shape extends earlier work on waves in the absence of shear and shows that the shape may be significantly altered by shear, the second-harmonic terms which describe the wave profile changing sign when the shear is increased sufficiently in an appropriate sense.

In the second part of the paper we show that the slope of internal waves at which breaking occurs (the particle speeds exceeding the phase speed of the waves) may be considerably reduced by the presence of shear. Internal waves on a thermocline which encounter an increasing shear, perhaps because of wind action accelerating the upper mixing layer of the ocean, may be prone to such breaking.

This work may alternatively be regarded as a study of the stability of a parallel stratified shear flow in the presence of a particular finite disturbance which corresponds to internal gravity waves propagating horizontally in the plane of the flow.

1. Introduction

The seasonal thermocline in the ocean and its cousin, the inversion which caps the atmospheric boundary layer, are regions of large vertical density gradient where the horizontal speed of the respective fluid often changes rapidly with depth or height. The change in speed across these regions is indeed often comparable with the speed at which internal gravity waves propagate through a still environment, and the shear may therefore be expected to produce significant effects on the phase speed of the waves, and perhaps also on their shape and on the conditions in which they break.

Few studies have been made of the effects of finite amplitude on internal waves. Benjamin (1966, who also gives comprehensive reference to other studies) found that nonlinear effects are important in those regions, like the thermocline, where the vertical density gradient is large. The present work was first stimulated by the observation that steady progressive interfacial waves in a two-layer fluid have *exactly* the same shape as have surface waves on a fluid of depth equal to the lower layer, provided they move at the speed of the upper layer (Thorpe 1974, appendix C). It is obviously necessary for the condition to be satisfied that the two layers are not

moving with the same speed). This result is in contrast to an earlier result that the shape of interfacial waves in a two-layer fluid with no shear is determined by the ratio of the depths and densities (Thorpe 1968*a*, hereafter called I). The waves then have narrow troughs and wide crests, the opposite of surface waves, if the density difference is small and the upper-layer depth much smaller than that of the lower layer.

The presence of shear in a two-layer fluid may thus change the wave shape, and a theory summarizing these changes is reviewed in appendix A. The two-layer fluid, whilst giving insight into the nature of the problem, is subject to the conceptual difficulty that, in the absence of surface tension at the interface and viscosity in the fluids, any shear flow is unstable, in the sense that there exist infinitesimal waves which grow exponentially† (Chandrasekhar 1961, chap. 11). This objection is avoided here by focusing attention on fluids in which the density and velocity vary continuously and have continuous derivatives, and which are thus closer to the fluids encountered in geophysical applications.

Banks, Drazin & Zaturska (1976) have given an account of small amplitude internal waves in shear flows, classifying the waves according to their stability properties. We shall here be concerned with the discrete class of modified internal gravity waves which (at least at small amplitude) are non-singular stable modes in a fluid with density increasing everywhere with depth. These waves have phase speeds c which lie outside the range of speeds of the mean flow (that is $c > \max U(z)$ or $c < \min U(z)$, where $U(z)$ is the mean flow speed and z is the upward vertical co-ordinate). The flows considered are two-dimensional with waves moving in the $+x$ direction at speeds which henceforth, without loss of generality, are supposed to exceed $\max U(z)$. In the limit of zero shear (or infinite Richardson number) the classical internal gravity waves are recovered. We shall not be concerned with the unstable wave modes which are associated with flows having small Richardson numbers and which coexist with the internal gravity waves in some parameter range but, recognizing that their presence implies that the flow is unstable, we shall draw some conclusions about the alternative mechanisms for instability when we come to discuss the stability of the waves and the experimental observations in §§ 5 and 6.

Orlanski & Bryan (1969) have shown that a change in the form of the waves, or breaking, occurs in internal waves (or in any wave system) when the forward speed of particles somewhere in the flow exceeds the phase speed of the waves. The physical transition accompanying such breaking in a stably stratified fluid leads inevitably to a region in which the density decreases with depth and in which, provided diffusion is negligible and the region can be maintained for a sufficient time, Rayleigh–Taylor instability and irreversible mixing processes may be expected. The subsequent physical flow may however again be stable, although it may contain 'rotors', regions of closed streamlines moving at the phase speed of the waves. Banks *et al.* (1976) showed that, as the flow Richardson number falls towards zero, the phase speed of small waves tends to the maximum speed of the mean flow. Thus for sufficiently small Richardson numbers, the wave motion need produce only a small increase to the forward speed

† This instability is correctly referred to as the Kelvin–Helmholtz instability, for these two scientists were concerned with the stability of sharp fluid interfaces. Recent practice however has been to use the name Kelvin–Helmholtz instability to describe the growth of small disturbances in *any* steady stably stratified shear flow, and we shall reluctantly continue this practice here.

of the flow to reach a condition of breaking at some level in the fluid. This suggested that a study of small, but finite, internal waves in shear might provide useful predictions of the onset of wave breaking. This theoretical work, and the associated experiments, may be regarded either as a study of finite amplitude waves and their stability in shear flow, or alternatively as a study of the stability of the flow itself to finite disturbances of a certain kind, namely those which resemble a uniform train of free waves. It should be remembered that the breaking instability is confined to a region which may be small in comparison with the waves and that, unlike Kelvin–Helmholtz instability, for example, the waves themselves do not grow in amplitude as the result of the instability, a feature which nicely distinguishes in a physical way the different types of instability. We shall find that for a given shear in a fluid of specified density profile (or for a given flow Richardson number in a specified flow) there is a minimum (or critical) wave slope above which waves of a given wavelength are unstable. This work also contrasts with the study of the stability of the wave train itself by a Benjamin–Feir-type mechanism (Borisenko *et al.* 1976; Grimshaw 1976; Thorpe 1977).

Our first concern is with the wave shape. The analysis is described in § 2.1. In the absence of analytical solutions, numerical solutions have been obtained for waves in fluids with density and velocity having hyperbolic-tangent forms (§ 3) and the results are compared with laboratory experiments (§ 4). Some results on the group velocity and the effect of finite amplitude on the phase speed are given in §§ 2.2 and 2.3 and the latter proves useful when we consider the breaking of the waves in §§ 5 (theory) and 6 (experiments). The theory is developed in conventional successive approximation form, following I, as far as third order, and although no formal convergence is established and the points of close comparison are few, the theoretical and experimental results are in general agreement.

Some remarks about the application of the results to natural flows are made in the concluding section, § 7.

2. Theory

2.1. *The wave shape*

We consider a stably stratified inviscid fluid of density $\rho_0(z)$ bounded between horizontal planes $z = 0$ and $z = h$, with z increasing upwards. We make the Boussinesq approximation† and suppose that the motion is two-dimensional, so that we may define a stream function $\psi(x, z, t)$ such that the velocity $\mathbf{u} = (\partial\psi/\partial z,\ 0,\ -\partial\psi/\partial x)$. The motion is governed by the vorticity equation

$$\rho_0(0)\,[\partial(\nabla^2\psi)/\partial t + J(\nabla^2\psi, \psi)] = g\,\partial\rho/\partial x, \tag{1}$$

where J is the Jacobian with respect to x and z, and ρ is the density of the fluid. The continuity equation is

$$\partial\rho/\partial t + J(\rho, \psi) = 0. \tag{2}$$

We suppose that at time $t = 0$ a progressive internal wave of wavenumber k is established in the horizontally infinite layer without, by horizontal inflow, changing the density structure, and that the mean Eulerian velocity of the fluid is $U(z)$. We

† This approximation may not always be valid for long waves. See, for example, Benjamin (1966).

look for a solution at subsequent times in the form

$$\psi = \sum_{n=-\infty}^{\infty} \phi_n E^n, \quad \rho = \sum_{n=-\infty}^{\infty} \rho_n E^n, \tag{3}$$

where $E = \exp[i(kx - \sigma t)]$, $\phi_{-n} = \tilde{\phi}_n$, $\rho_{-n} = \tilde{\rho}_n$ (a tilde denotes the complex conjugate) and σ is the wave frequency (yet to be determined). We also write

$$\phi_n = \sum_{j=n}^{\infty} \epsilon^j \phi_{nj}, \quad \rho_n = \sum_{j=n}^{\infty} \epsilon^j \rho_{nj}, \tag{4}$$

where ϕ_{nj} and ρ_{nj} are functions of z, ξ and τ only and ϵ is an expansion parameter. Here

$$\xi = \epsilon(x - c_g t), \quad \tau = \epsilon^2 t \tag{5}$$

are slowly varying scales, and c_g is the horizontal group velocity of the waves (which is also to be determined; the expansion technique follows that of Davey & Stewartson (1974), being common to a number of multiple-scale approaches). We suppose that

$$\rho_{00} = \rho_0(z), \quad \phi_{00,z} = U(z). \tag{6}$$

The expressions for ψ and ρ are now substituted into (1) and (2) and the method of multiple scales used to obtain a series of differential equations for the ϕ_{nj} and ρ_{nj}. These are solved by successively equating the coefficients of $\epsilon^j E^n$ to zero and using (2) to remove the ρ_{nj} from (1) to obtain ordinary differential equations for the ϕ_{nj} in terms of known ρ's and ϕ's with smaller suffix values. These equations are solved for ϕ_{nj} and the equation derived from (2) used to find ρ_{nj}. The equation of a constant-density surface $z + \eta(x, z, t, \xi, \tau)$ may be found by expressing η as a double series in ϵ and E similar to (3) and (4), expanding $\rho(x, z+\eta, t, \xi, \tau) - \rho_0(z) = 0$ as a Taylor series and comparing coefficients to find the η_{nj}. We impose the condition that the mean displacement shall be zero so that the disturbed fluid has the same mean density distribution as the undisturbed fluid. This condition has been described in detail elsewhere (I; Yih 1974) and is equivalent to the condition that the mean fluid depth is unchanged in the study of finite amplitude surface waves. The phase speed of the infinitesimal waves is determined by comparing coefficients of ϵE, the group velocity by comparing coefficients of $\epsilon^2 E$ and the effect of finite amplitude on the phase speed by comparing coefficients of $\epsilon^3 E$.

On comparing coefficients of ϵE we find

$$\phi_{11} = A\Psi(z), \quad \rho_{11} = \rho_0' A\Psi/(U-c), \quad \eta_{11} = -A\Psi/(U-c), \tag{7}$$

where $A = A(\xi, \tau)$ is an undetermined function of the slowly varying scales, and

$$\mathscr{L}(k, c)\Psi \equiv \left[\frac{d^2}{dz^2} - \left(k^2 - \frac{N^2}{(U-c)^2} + \frac{U''}{U-c}\right)\right]\Psi = 0 \tag{8}$$

with $\Psi(0) = \Psi(h) = 0$. This is the Taylor–Goldstein equation (Thorpe 1969) familiar in instability problems; c is the phase speed of the waves and N is the Brunt–Väisälä frequency. N^2 is nowhere negative since the fluid is stably stratified, and N is therefore real. We assume that, given k, there are solutions of (8) which satisfy the boundary conditions with a set of real eigenvalues of c which lie outside the range of U, so that $U - c$ is everywhere non-zero (Banks *et al.* 1976). These represent internal gravity wave modes and of these we select one eigensolution $\Psi(z)$ and the corresponding

phase speed c. At order ϵ^2E^2 we find terms which contribute to the wave shape. The stream function is given by $\phi_{22} = A^2\Psi_2(z)$, where†

$$\mathscr{L}(2k,c)\Psi_2 = \frac{\Psi^2}{2(U-c)}\left\{U''' + \frac{2g\rho_0''}{\rho_0(0)(U-c)} - \frac{U'U''}{U-c} + \frac{3N^2U'}{(U-c)^2}\right\}, \tag{9}$$

whilst

$$\rho_{22} = A^2\left\{\frac{\rho_0'\Psi_2}{U-c} + \frac{\Psi^2}{2(U-c)}\left(\rho_0'' - \frac{\rho_0'U'}{U-c}\right)\right\} \tag{10}$$

and

$$\eta_{22} = A^2\left\{\frac{\Psi\Psi'}{(U-c)^2} - \frac{\Psi_2}{U-c} - \frac{\Psi^2U'}{2(U-c)^3}\right\}. \tag{11}$$

For a uniform train of waves the wave 'shape' is given locally, to order ϵ^2, by

$$\begin{aligned}\eta &= \epsilon\eta_{11}E + \epsilon^2\eta_{22}E^2 + \text{complex conjugate}\\ &= \frac{2\epsilon|A|\Psi}{c-U}\cos(kx-\sigma t+\delta) + 2(\epsilon|A|)^2\left\{\frac{\Psi\Psi'}{(c-U)^2} + \frac{\Psi_2}{c-U} + \frac{\Psi^2U'}{2(c-U)^3}\right\}\cos 2(kx-\sigma t+\delta)\\ &= \epsilon\eta_1(z)\cos(kx-\sigma t+\delta) + \epsilon^2\eta_2(z)\cos 2(kx-\sigma t+\delta),\end{aligned} \tag{12}$$

say, where $\delta = \arg A$ (other terms in ϵ and ϵ^2 are zero, as we show below). For those values of z for which η_2 is positive the isopycnal surfaces have a wave form with relatively narrow crests and wide troughs, whilst if η_2 is negative the reverse is true. This result reduces to that found in I when $U' = 0$.

2.2. *The group velocity*

Comparing coefficients of order ϵE^0, ϕ_{01} and ρ_{01} are found to be zero when the conditions on the mean flow and mean displacement are applied. Hence

$$\phi_{01} = 0, \quad \rho_{01} = 0, \quad \eta_{01} = 0. \tag{13}$$

At order ϵ^2E^1 the effects of an x variation in A first become significant. The equation for ϕ_{12} is

$$\mathscr{L}(k,c)\phi_{12} = \frac{iA_\xi}{k}\Psi\left\{\frac{(c_g-c)}{(U-c)^3}[2N^2 - U''(U-c)] - 2k^2\right\}. \tag{14}$$

The operator $\mathscr{L}(k,c)$ is self-adjoint and the left-hand side of the equation is identical to that of (8), and so following Stuart (1960), solutions of (14) exist, subject to the boundary conditions $\phi_{12} = 0$ at $z = 0$, h, if and only if the following orthogonality condition holds:

$$\int_0^h \Psi^2\left\{\frac{c_g-c}{(U-c)^3}[2N^2 - U''(U-c)] - 2k^2\right\}dz = 0. \tag{15}$$

This may be reorganized to give

$$c_g = c\left[1 - 2\sigma^2\int_0^h \Psi^2 dz \Big/ c^3\int_0^h \frac{\Psi^2}{(c-U)^3}[2N^2 + U''(c-U)]\,dz\right]. \tag{16}$$

When $U = 0$ this gives the more familiar equation for the group velocity of waves in the absence of shear (Thorpe 1977). Since

$$c_g = \frac{\partial\sigma}{\partial k} = \frac{\partial}{\partial k}(kc) = c + k\frac{\partial c}{\partial\alpha}$$

† Note that, in contrast to I (equation 3.2.19), there is now no complete orthogonal set of eigenfunctions which can be used to express the left-hand side of (9).

it follows from (16) that

$$\frac{\partial c}{\partial \alpha} = -2k \int_0^h \Psi^2 dz \Big/ \int_0^h \frac{\Psi^2}{(c-U)^3}[2N^2 + U''(c-U)]\,dz, \tag{17}$$

a result found by perturbation techniques by Howard (1963, equation 13) in his study of unstable waves near stability boundaries, and recently by Voronovich (1976) in rather more general conditions than those considered here. Since c lies outside the range of U, $c-U$ is positive and so, if $U'' > 0$ everywhere, it follows that $c_g \leqslant c$. The group velocity c_g may however lie within the range of U, and an example with this property is discussed in § 3.

2.3. *Finite amplitude effects on phase speed*

Having considered in the last subsection linear waves subjected to a modulation which moves at the group velocity, we now turn to the effect of finite amplitude on the phase speed of a train of waves which is uniform in the x direction, in which there is no ξ variation.

A solution of (14) is now $\phi_{12} = 0$ and it follows that

$$\rho_{12} = 0, \quad \eta_{12} = 0. \tag{18}$$

Comparing coefficients at order $\epsilon^2 E^0$ yields $\phi_{02} = 0$, and the condition of zero displacement gives $\eta_{02} = 0$, from which it follows that

$$\rho_{02} = \frac{|A|^2}{4}\left(\frac{\rho_0' \Psi^2}{(U-c)^2}\right)'.$$

The correction to the phase speed is found at order $\epsilon^3 E$, where the governing equation is

$$\mathscr{L}(k,c)\,\phi_{13} = \frac{2iA_\tau \Psi}{k(c-U)^3}[2N^2 + U''(c-U)] + \chi(z)|A|^2 A, \tag{19}$$

where $\chi(z)$ is a determined function of z.

As before, a solution satisfying the boundary conditions can be found only if an orthogonality condition holds:

$$A_\tau \int_0^h \frac{\Psi^2}{k(c-U)^3}[2N^2 + U''(c-U)]\,dz = -iA|A|^2 \int_0^h \chi(z)\Psi\,dz. \tag{20}$$

This is an equation for A which has the solution

$$A = a \exp(-i\sigma_2 a^2 \tau), \tag{21}$$

where σ_2 is given in appendix B and $a = |A|$ is real. Substitution in (12) gives a wave with phase $kx - (\sigma + \epsilon^2 a^2 \sigma_2)t$, which thus advances with speed $(\sigma + \epsilon^2 a^2 \sigma_2)/k$. The effect of finite amplitude is thus to increase the phase speed by a fraction

$$\epsilon^2 c_2 / c = \epsilon^2 |A|^2 \sigma_2 / \sigma.$$

Because of the occurrence of high-order derivatives in the expression for σ_2, it seems unlikely that any general conclusions about the sign of c_2 can easily be deduced.

It will later be found useful to have an expression for ϕ_{33} in the case when A is independent of ξ. This has been calculated by comparing coefficients of $\epsilon^3 E^3$. We find

$\phi_{33} = A^3\Psi_3(z)$, where $\Psi_3(z)$ is given in appendix B. We can then write, correct to third order,

$$\psi = \Psi_0(z) + \epsilon\psi_1(z)\cos\zeta + \epsilon^2\psi_2(z)\cos 2\zeta + \epsilon^3\psi_3(z)\cos 3\zeta, \tag{22}$$

where $\psi_1 = 2|A|\Psi(z)$, $\psi_2 = 2|A|^2\Psi_2(z)$, $\psi_3 = 2|A|^3\Psi_3(z)$ and $\zeta = k[x-(c+\epsilon^2 c_2)t]$. At this stage we can of course incorporate the parameter ϵ with A, or equivalently, put $\epsilon = 1$.

3. Numerical examples

Even in the few cases† in which exact analytical solutions of (8) can be found which are continuous in a range in which the flow Richardson number varies from infinity to some finite value, the task of solving (9) or of performing the integrals in (16) is considerable, and we have therefore investigated numerically some profiles which not only represent those found in the natural environment but which can approximately be reproduced in laboratory experiments.

We have examined the profiles

$$U = U_0\tanh\beta(z-d), \quad \rho_0(z) = \rho_0(0)\,[1-\Delta\tanh\beta(z-d)], \tag{23}$$

which have inflexion points at $z = d$, $0 < d < h$, and an interfacial thickness $2\beta^{-1}$. It should be noticed that if U_0 is positive there is a flow for $z < d$ which opposes the direction of wave propagation. We looked for solutions of (8) which correspond to the first wave mode for which $\eta_1(z)$ is zero at $z = 0$, h only, using a shooting method with 200 points in the vertical and with integrations performed using Simpson's rule.

Figure 1 shows the variation of c/c_0 and c_g/c_0 with $U_0/c_0 \equiv \delta$, say, where c_0 is the value taken by c when U_0 is zero, for $d = 0{\cdot}25h$, $\beta h = 20$ and $kh = \pi$. The numerical solution gave $c_0 = 0{\cdot}457(g\Delta h)^{\frac{1}{2}}$. As expected, c is everywhere positive, lying outside the range of U. However c_g takes values within the range of U for approximately

$$-1{\cdot}3 < \delta < -0{\cdot}6.$$

The minimum gradient Richardson number of the mean flow is found at $z = d$ and is

$$Ri = g\Delta/\beta U_0^2 = 0{\cdot}239/\delta^2.$$

By the Miles–Howard theorem, small disturbances will be stable provided that $Ri > \frac{1}{4}$, or $|\delta| < 0{\cdot}978$. In the absence of boundaries this flow would be unstable for $Ri < \frac{1}{4}$, and it is probable that this condition continues to apply here (see Hazel 1972, figure 3. Hazel's parameter Y is roughly equivalent to βd, which is 5 in the present case).

The internal gravity wave solutions continue into the regions in which unstable waves coexist. At small Ri, both c and c_g tend to U_0, as found by Banks *et al.* (1976). Normalized profiles of N^2, U/c_0, ψ_1, ψ_2, ψ_3, η_1 and η_2 for various values of δ are shown in figure 2. The normalization is such that the maximum values (except for U/c_0) are unity. The table given in the caption to figure 2 shows how nonlinear effects increase in importance as $|\delta|$ increases. The level of the maximum value of ψ_1 descends as δ increases, whilst the maximum of η_1 rises in level from well below the interface at $z = 0{\cdot}25h$ to well above it. The change in sign of η_2 between $\delta = -0{\cdot}311$ and $-0{\cdot}933$ results in a change in the shape of the wave [see (12)]. At $\delta = -0{\cdot}62$ the shape of the

† For example see that of Miles (1967) or Taylor (1931) and Eliassen, Høiland & Riis (1953).

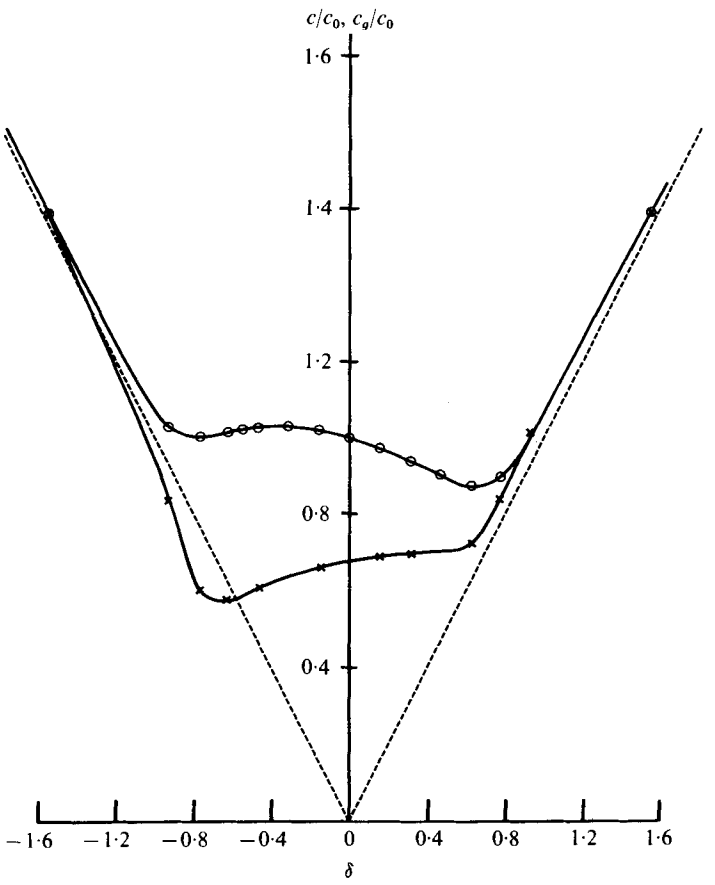


FIGURE 1. The variation of the phase speed c and group velocity c_g with U_0, non-dimensionalized with respect to the phase speed c_0 at $U_0 = 0$, for the density and velocity profiles given by (23) with $d = 0{\cdot}25h$, $\beta h = 20$, $kh = \pi$.

isopycnal surfaces depends on depth, taking both positive and negative values, and both narrow and wide crests are present. These results are shown more clearly in figure 3 (a), where the positions of the maximum and minimum values of η_1 and η_2 are shown. Figure 3 (b) shows the same variables for a fluid with a narrower interface, $\beta h = 30$. These results may be compared with those in the equivalent two-layer fluid, where the change in shape occurs at $\delta = -0{\cdot}26$ (see appendix A), and a trend in the

FIGURE 2. Normalized profiles of N^2, U/c_0, ψ_1, ψ_2, ψ_3 [see (22)], η_1 and η_2 [see (12)] for various δ with $kh = \pi$, $\beta h = 20$, $d = 0{\cdot}25h$. The maximum values of ψ_2, ψ_3, η_1 and η_2 are as follows when the maximum value of ψ_1 is unity:

δ	max ψ_2	max ψ_3	max η_1	max η_2
−0·93	$2{\cdot}89 \times 10^{-2}$	$5{\cdot}32 \times 10^{-3}$	3·21	$3{\cdot}27 \times 10^{-1}$
−0·62	$1{\cdot}26 \times 10^{-2}$	$1{\cdot}91 \times 10^{-4}$	1·38	$7{\cdot}78 \times 10^{-3}$
−0·31	$7{\cdot}52 \times 10^{-3}$	$3{\cdot}73 \times 10^{-4}$	1·39	$8{\cdot}41 \times 10^{-3}$
0·0	$1{\cdot}65 \times 10^{-2}$	$4{\cdot}67 \times 10^{-4}$	1·56	$2{\cdot}18 \times 10^{-2}$
0·31	$2{\cdot}94 \times 10^{-2}$	$8{\cdot}58 \times 10^{-4}$	1·74	$4{\cdot}76 \times 10^{-2}$
0·62	$5{\cdot}57 \times 10^{-2}$	$4{\cdot}13 \times 10^{-3}$	2·56	$1{\cdot}84 \times 10^{-1}$
0·93	1·75	4·24	16·8	35·2

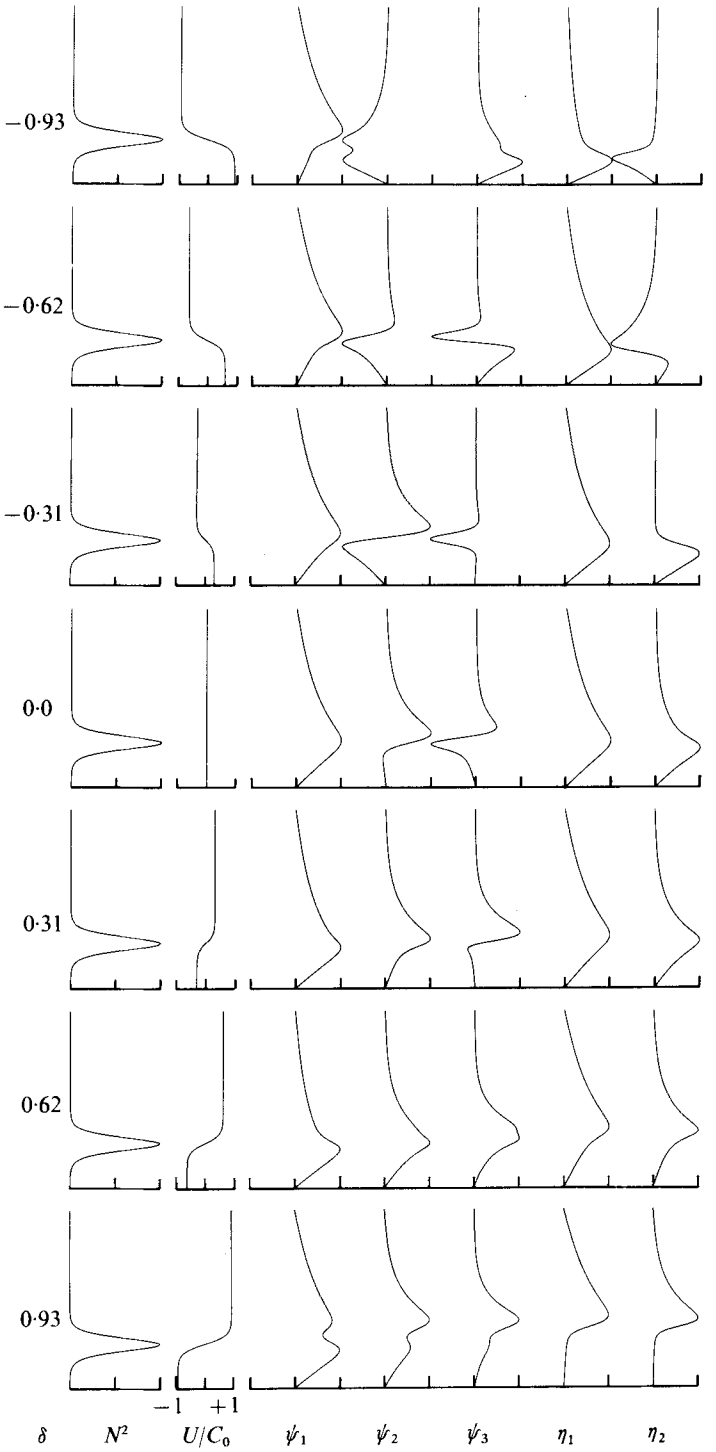


FIGURE 2. For legend see facing page.

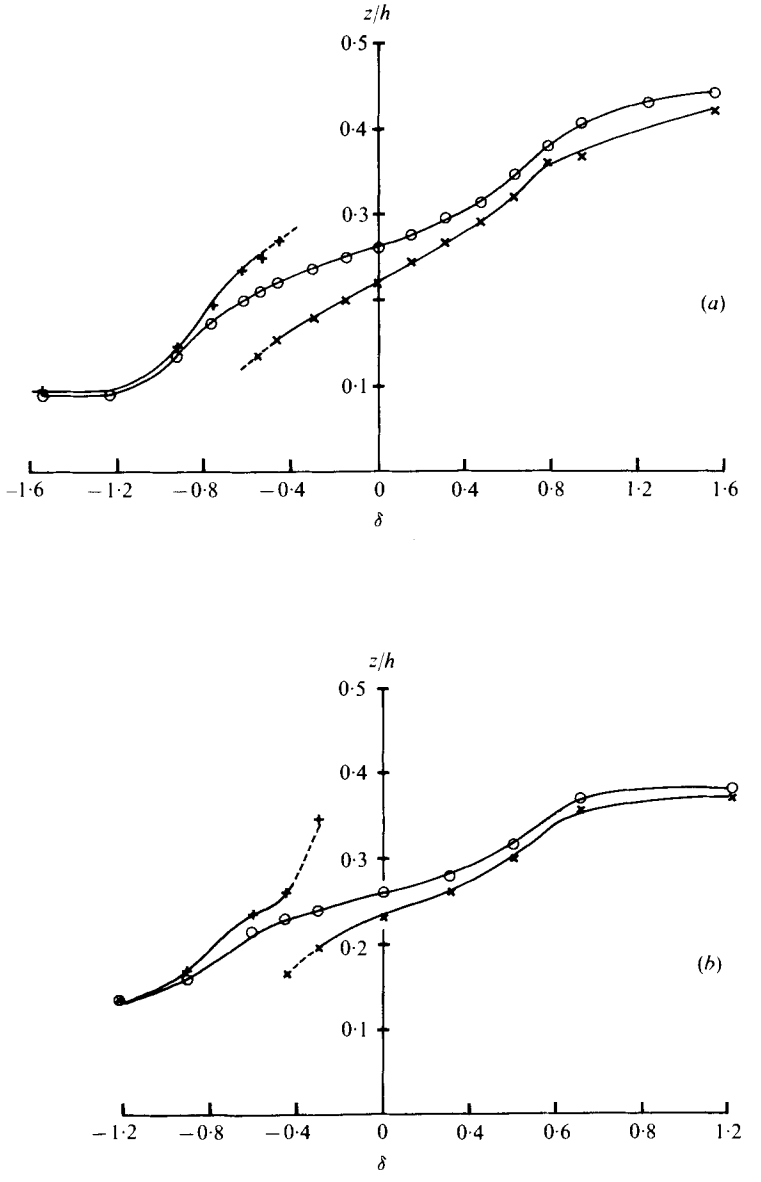


FIGURE 3. The variation with δ of the depths z/h at which η_1 and η_2 have maximum or minimum values other than at $z = 0$ or h for waves with $kh = \pi$, $d = 0{\cdot}25h$ and (*a*) $\beta h = 20$, (*b*) $\beta h = 30$. ⊙, maximum of η_1; ×, maximum of η_2; +, minimum of η_2. Where the curve is dashed the magnitude of η_2 is not the largest which exists in the flow.

direction of this value as the interface becomes narrower may be seen in figures 3(*a*) and (*b*). The displacement of the maximum of η_1 (and ψ_1) from the level of the density inflexion for large $|\delta|$, however, suggests that comparison with the two-layer model is bound to be of limited scope and that the qualitative similarities with surface waves noted in § 1 may not be found when the density interface is of finite thickness. In particular a parameter which is a measure of the wave shape is $\max \eta_2/c(k \max \eta_1)^2$,

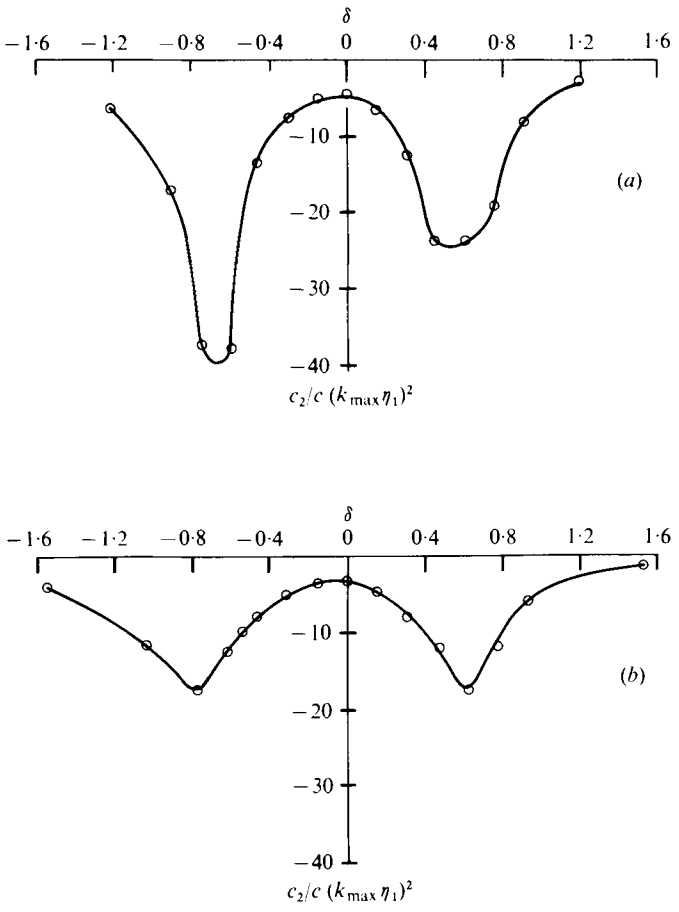


FIGURE 4. The variation of $c_2/c(k \max \eta_1)^2$ with δ for $kh = \pi$, $d = 0{\cdot}25h$ and (*a*) $\beta h = 20$, (*b*) $\beta h = 30$.

and for surface waves in water of depth h this is equal to $\frac{1}{4}(3\tanh^{-3}kh - \tanh^{-1}kh)$ (Stokes 1947). When $|\delta|$ is large, so that c is close to U_0, a similarity between the estimates of the slope parameters of the internal waves with $d = \frac{1}{4}h$ and of surface waves with $kh = \frac{1}{4}\pi$ ($\delta > 0$) or $kh = \frac{3}{4}\pi$ ($\delta < 0$) might be expected, since then the corresponding interfacial waves ($\beta h \to \infty$) are known to be identical in shape to the surface waves. In practice, however, the estimates for the waves shown in figure 3 are much in excess of those of the corresponding surface waves.

Figure 4 shows the variation of $c_2/c(k \max \eta_1)^2$ with δ, where $k \max \eta_1$ is a measure of the slope of the isopycnal surfaces. Unlike surface waves, for which $c_2 > 0$, c_2 is negative for the waves considered here, although it is known that internal waves with $c_2 > 0$ do occur (Thorpe 1977).

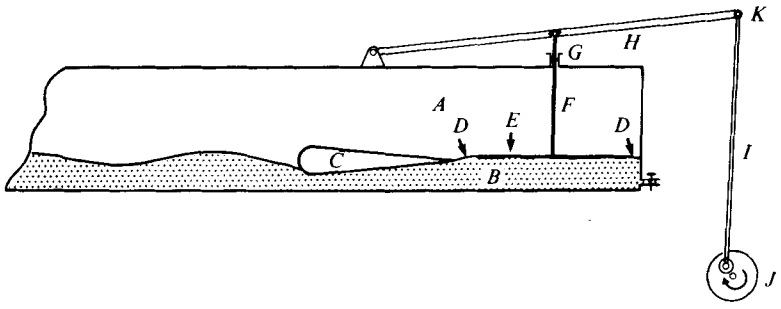


FIGURE 5. The wave maker. A is the almost fresh water, B the brine solution, C is a fixed 'aerofoil' shape, spanning the tube, and D is a rubber diaphragm connecting the horizontal plate E to the tube walls and separating the water at A from the brine at B. The plunger F is driven up and down via the linkage rods H and I from the rotating disk J turned by an electric motor via a hydraulic gearbox. As the plate E is raised water is expelled from the region A and drawn into B, causing a depression in the interface. A downward motion reverses the process. Repeated oscillations drive a train of waves. When the tube is to be tilted, the linkage pin K is drawn out, releasing the rod I and disconnecting the tube from the drive.

4. Experiments on the wave shape

We have made some simple laboratory experiments to test the theoretical and numerical results on the wave shape.

The apparatus consists of a 4·85 m rectangular tube with Perspex side walls which is of height 16 cm and width 10 cm and which may be rotated about a horizontal axis. A wave maker similar to that used by Keulegan & Carpenter (1961), and shown in figure 5, is at one end of the tube. The tube is filled in an inclined position, first with water, in which a little salt has been dissolved to aid conductivity measurements, and then with a denser brine solution (sometimes coloured with potassium permanganate), until the air has been completely removed and the brine–water interface is one-quarter of the way up the tube. It is then carefully rotated into a horizontal position, the brine spreading under the water and eventually forming a layer, 4 cm in depth, having a sharp interface with the water. The interface is allowed to settle and diffuse. Vertical profiles are made with a single-electrode conductivity probe, from which the density can be inferred, until the desired density profile is obtained.† Waves of the first internal mode are then produced by the wave maker. If the piston motion is too great the waves spill over the solid 'aerofoil' section, water being drawn into the region B beneath the diaphragm or brine into region A above, and mixing ensues. Care was taken to avoid such conditions in the experiments. As with most wave makers the wave form is not ideal, and in particular a small free second-harmonic component is generated which distorts the primary wave form.

The waves are made visible by shadowgraphs (e.g. figures 6 and 12, plates 1, 8 and 9) or with dyes (e.g. figures 9 and 10, plates 2–5) or by the distortions they produce when a set of parallel lines is viewed through the tube (e.g. figure 11, plates 6 and 7). Before the waves have reached the far end of the tube, the tube is tilted through a small angle to induce a shear flow, and then returned to the horizontal once the flow

† The density profile is perhaps closer to a diffusion-produced error-function profile, but for convenience we approximate it by a hyperbolic-tangent profile in the calculations.

is established. Photographs are taken with a 16 mm ciné (23 frames s^{-1}) and a 35 mm 'still' camera (taking about 2 frames s^{-1}) to record the changes in wave form.

The flow generated in a tilted tube has been examined by Thorpe (1968*b*). If the density distribution is $\rho(z)$ the flow parallel to the axis of the tube in the absence of waves has the form

$$u(z,t) = gt\sin\alpha\left(\frac{h}{\rho\int_0^h \frac{dz}{\rho} - 1}\right),$$

when viscosity can be neglected, where α is the angle of tilt and h is the tube depth. (The effect of viscosity has been considered by Thorpe (1971, appendix). The parameter Q which appears there in equation (A 5) is at the smallest 0·87 in the present experiments, and we henceforth neglect viscosity.) If $\rho = \rho_0[1 - \Delta\tanh\beta(z-d)]$ and $\Delta \ll 1$,

$$u(z,t) = g\Delta t\sin\alpha[\tanh\beta(z-d) - (h-2d)/h].$$

This flow is maintained if, at time t, the tube is brought into a horizontal position, at least unless Kelvin–Helmholtz instability occurs or until the arrival of surges from the ends of the tube.

A study of the interaction between waves and accelerating shear flow has been attempted only for small amplitude waves in a two-layer fluid (Thorpe 1969). It was shown that the waves developed as parabolic cylinder functions and that, provided the angle of tilt was sufficiently small, a quasi-steady approximation gave a good description of the wave development. The analysis of finite waves will not be attempted here and we shall proceed on the assumption that the mean zero-order flow after the tube has been tilted for a time t_1 is given by

$$U(z) = U_0[\tanh\beta(z-d) - \tfrac{1}{2}], \tag{24}$$

where $U_0 = g\Delta t_1\sin\alpha$ and we have used $d = \frac{1}{4}h$. The interface $z = d$ is advected with speed $-\frac{1}{2}U_0$ and thus the phase and group speeds are reduced (or increased if α, and therefore U_0, is negative).

An interface thickness $2\beta^{-1}$ of $1{\cdot}01 \pm 0{\cdot}08$ cm (one standard deviation is given) was used in these experiments, which gives a value $\beta h = 31{\cdot}7 \pm 2{\cdot}2$. When broader interfaces were tried the operation of the wave maker was less than satisfactory, there being more mixing at the diffuse 'edges' of the interface, and it was for this reason that a rather narrow interface was used. The density difference Δ was $(25{\cdot}3 \pm 0{\cdot}5) \times 10^{-3}$ and the wave frequency was $1{\cdot}83 \pm 0{\cdot}02$ rad s^{-1}, which resulted in a wavenumber such that $kh = 3{\cdot}19 \pm 0{\cdot}02$. These values are close to those examined numerically (figures 3*b*, 4*b*). The waves before the tube was tilted ($\delta = 0$) were in each case produced by the same vertical motion of the piston driving the waves, and have a height-to-wavelength ratio of $0{\cdot}034 \pm 0{\cdot}003$. Their wavelength is some five times greater than that of the fastest growing Kelvin–Helmholtz waves in the parallel shear flow (Thorpe 1971), which have a length $4\pi/\beta$, or a wavenumber k_1 such that $k_1 h = \frac{1}{2}\beta h$, approximately 15·8.

The angle of tilt of the tube was $(5{\cdot}09 \pm 0{\cdot}03) \times 10^{-2}$ rad, so that $\delta = 0{\cdot}5$ after approximately 1·06 wave periods. Figure 6 (plate 1) shows shadowgraphs of the waves for a variety of values of δ approximately one period after the tube was returned to the horizontal position. The detailed changes in shape are obscured by the presence of the free second harmonic but the general trend follows the predicted pattern,

passing from waves with relatively narrow troughs at large negative δ to waves with markedly narrower crests at positive δ, but it is not possible to determine the precise value of δ at which the change occurs. Changes in shape might have been easier to detect if larger waves had been used but, as we shall show later (e.g. figure 13), these are prone to instability at smaller values of $|\delta|$ and it is impossible with much larger waves to reach the negative value of δ at which the second-harmonic contribution to the wave profile changes sign.

In the first experiments which were made, the tube was not returned to the horizontal but remained tilted so that $|\delta|$ increased linearly with time. The evolution of the wave shape followed the trend with δ which we have described but led eventually to a breakdown in the flow. That the shape of the waves is significantly affected by the direction of the shear is perhaps most persuasively demonstrated by figures 9 and 10 (plates 2–5), which show the waves before and during breaking at positive and negative δ respectively. Before discussing these experiments further we shall consider the conditions which lead to the breaking of internal waves.

5. Breaking internal waves

5.1. *Particle orbits*

Internal waves will begin to break when the forward particle speed at some point in the fluid exceeds the phase speed† or when Kelvin–Helmholtz instability is induced by the wave motion. We here consider the first possibility when a shear flow accompanies the waves.

Since the horizontal particle speed is $\partial\psi/\partial z$ we may immediately proceed to search for the region in the flow at which, for a given ϵ, the forward speed is largest, and then, by varying ϵ, find the smallest value of ϵ, and thus the wave slope, at which this speed just equals the phase speed. It is important however, particularly when the effect of the instability on the density distribution is considered, to know not only where in the flow the instability occurs, but at what density level it occurs, overturning in the region of large vertical gradient having a greater potential effect on the density profile and energy of the fluid than overturning at the edges. We shall therefore consider the motion of particles and develop a description of the flow in terms of the particle orbits. The motion of the particles is found about the mean orbital level z_0, which therefore specifies the particle density.

Relative to a stationary frame of reference, the motion of a particle is given by co-ordinates (x, z), where

$$\left.\begin{aligned} x &= \Psi_0'(z_0)\,t + x_0 + \epsilon x_1 + \epsilon^2 x_2 + \ldots, \\ z &= z_0 + \epsilon z_1 + \epsilon^2 z_2 + \ldots, \end{aligned}\right\} \tag{25}$$

say, where x_1, x_2, z_1, $z_2 \ldots$ describe the displacements from a position (x_0, z_0) moving with the basic speed $U(z_0)$. Now if A is real and independent of t, the (Eulerian) stream

† Consider the steady motion in a frame of reference moving forward at the phase speed of the waves. In the x, z plane lines of constant density are particle paths, the particles in general moving in the $-x$ direction. If the particle speed exceeds the phase speed at some point in the flow, the particle motion is reversed and the particle path, and hence the line of constant density passing through this point, must assume an S-shape, and hence some neighbouring fluid must be gravitationally unstable at the point.

function is given by (22) and the motion at a point is given by

$$dz/dt = -\partial\psi/\partial x, \quad dx/dt = \partial\psi/\partial z. \tag{26}$$

Substituting into (26) and collecting terms of order ϵ we find

$$dz_1/dt = k\psi_1(z_0)\sin\zeta_1, \quad \text{where} \quad \zeta_1 = k[x_0 + (U(z_0) - c)\,t],$$

and so

$$z_1 = \frac{\psi_1(z_0)}{c-U}\cos\zeta_1. \tag{27}$$

(The coefficient of $\cos\zeta_1$ is proportional to $\eta_1(z)$.) Also

$$\frac{dx_1}{dt} = \left(\frac{\psi_1\Psi_0''}{c-U} + \psi_1'\right)\cos\zeta_1 \tag{28}$$

(using 26) and so

$$x_1 = \frac{1}{k(U-c)}\left(\frac{\psi_1\Psi_0''}{c-U} + \psi_1'\right)\sin\zeta_1,$$

where ψ_1, U, Ψ_0'' etc. are evaluated at $z = z_0$. The horizontal particle speed dx_1/dt is made up of two contributions: the Eulerian part ψ_1' and a part which derives from the vertical motion of amplitude z_1 in the shear flow, the particle being carried backwards and forwards as it moves up and down in the relative shear flow.

Collecting and equating terms of order ϵ^2, we find

$$\frac{dz_2}{dt} = k\left[2\psi_2 - \frac{\psi_1^2\Psi_0''}{2(c-U)^2}\right]\sin 2\zeta_1$$

and so

$$z_2 = \frac{1}{2(c-U)}\left[2\psi_2 - \frac{\psi_1^2\Psi_0''}{2(c-U)^2}\right]\cos 2\zeta_1 \tag{29}$$

(there is zero mean vertical displacement at this order),† and

$$\frac{dx_2}{dt} = \theta_1(z_0)\cos 2\zeta_1 + \theta_2(z_0)$$

(where $\theta_1(z)$ and $\theta_2(z)$ are given in appendix B) and so

$$x_2 = \frac{\theta_1(z_0)}{2k(U-c)}\sin 2\zeta_1 + \theta_2(z_0)\,t. \tag{30}$$

Hence relative to $U(z_0)$ particles move forward with a speed $\theta_2(z_0)$. This is the second-order mean drift, already familiar for surface waves and internal gravity waves (I, appendix 6; Thorpe 1977).

At third order the finite amplitude correction to the phase speed enters the expression (22) for ψ. Recommencing the expansion we find, correct to third order, that the forward speed of the particles is

$$dx/dt = \Psi_0' + \epsilon^2\theta_2 + \epsilon(\psi_1' + \psi_1\Psi_0''/c_3 + \epsilon^2\theta_3)\cos\zeta_2 + \epsilon^2\theta_1\cos 2\zeta_1 + \epsilon^3\theta_4\cos 3\zeta_2, \tag{31}$$

† Coefficient of $\cos 2\zeta_1$ appearing in the particle displacement (29) is not equal to η_2, the coefficient at corresponding order in the equation for a line of constant density, since the two quantities are measured in different frames of reference. Equation (25) expresses the relation between the two frames. The difference is however negligible at order ϵ in η or the particle displacement, as noted under (27).

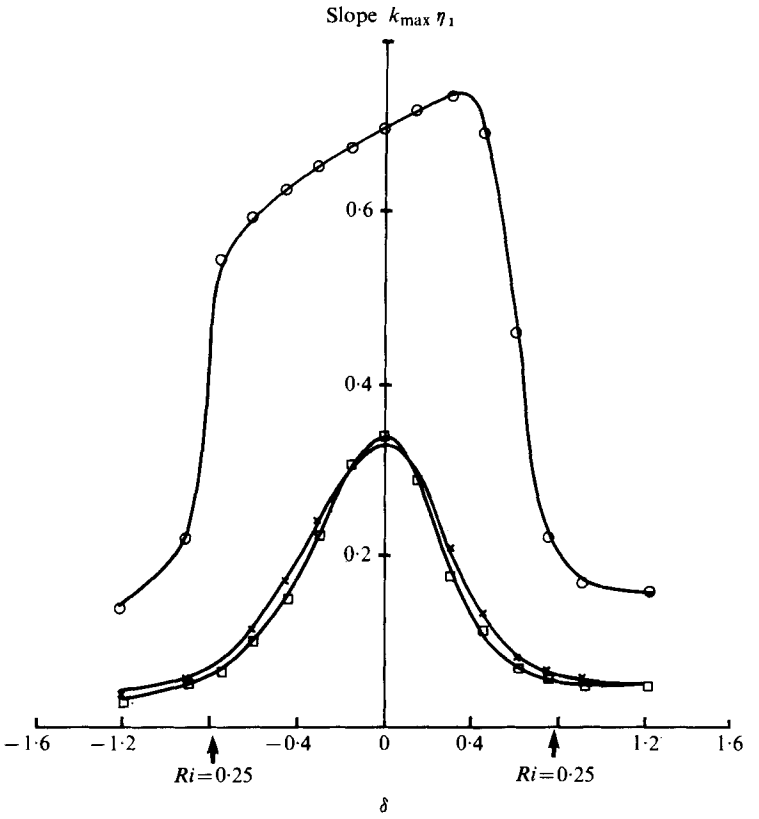


FIGURE 7. The smallest wave slope, when δ is varied, at which the horizontal velocity of particles somewhere in the flow is just equal to the phase speed of the waves. ⊙, points determined taking into account only first-order terms; ×, including second-order terms; ⊡, including the third-order terms as given by (31). The mean flow is given by (23) and $kh = \pi$, $\beta h = 30$, $d = 0{\cdot}25h$.

where $\zeta_2 = k[x_0 - c_3 t]$ and $c_3 = c - U + \epsilon^2(c_2 - \theta_2).\theta_3(z)$ and $\theta_4(z)$ are given in appendix B, and the z variables are evaluated at $z = z_0$.

This forward particle speed is to be compared with the phase speed $c + \epsilon^2 c_2$, which is correct to order ϵ^3.

5.2. *Numerical calculations*

The smallest wave slope ($k \max \eta_1$) for which the phase speeds and maximum horizontal particle speeds ($\max u$) are equal at orders ϵ, ϵ^2 and ϵ^3 have been calculated numerically for the tanh profiles (23) with $d = 0{\cdot}25h$, $\beta h = 30$ and $kh = \pi$. As before 200 points were examined in the vertical (z) direction but in the horizontal (or time) co-ordinate the slope was examined at phases of $\frac{1}{8}n\pi$, $n = 0, 1, 2, \ldots, 8$, zero phase corresponding to the wave crest and a phase π to the trough. The results are presented in figures 7 and 8. The inclusion of second-order terms reduces the wave slope at which, for a given value of δ, the waves become unstable to breaking, but the third-order terms produce only a small further reduction over most of the curve. (This result is not however true in general.) Figure 7 may be thought of as a stability diagram, with slopes lying below the 'critical' slopes on the curve in a given shear flow (that is

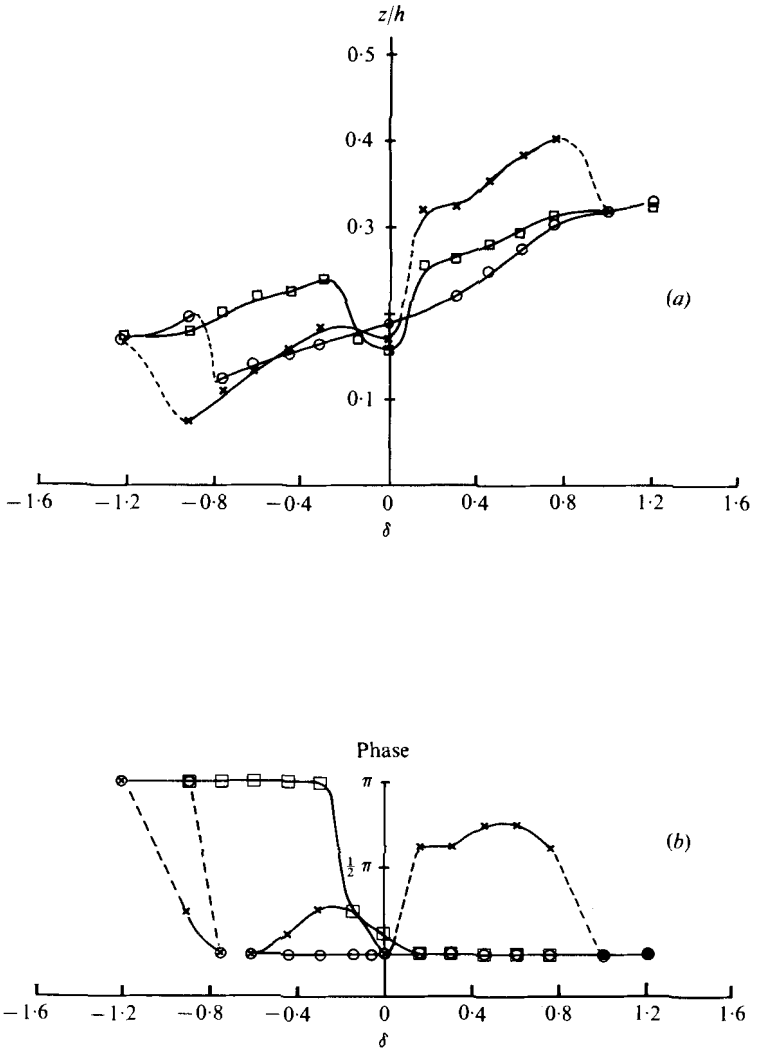


FIGURE 8. (*a*) The mean depth z/h of the particles and (*b*) the phase of the wave when the conditions specified in the caption to figure 7 are satisfied. Symbols as in figure 7.

at fixed δ) being stable, whilst those with high slopes are unstable. If only first-order effects are considered there is a region ($0 < \delta < 0{\cdot}4$) in which an increase in shear for waves with slopes lying just above the curve makes the waves stable, but for the second- and third-order solutions an increase in shear with a stable wave, or an increase in its slope, results in a trend towards an unstable situation. The presence of shear results in a considerable reduction in the slope at which waves break, and a corresponding large reduction in the wave energy. An asymmetry for positive and *negative* δ *is retained even at third order, a greater* value of $|U_0|$ being needed to cause breaking at $\delta < 0$ than for waves at the same slope (in the range $0 < |\delta| < 0{\cdot}9$) for $\delta > 0$, a feature which we shall find reflected in the experimental results. The critical wave slopes determined directly from $u = \partial\psi/\partial z$ show good agreement with those found from the particle orbits.

The asymmetry of the first-order curve near $\delta = 0$ may be explained as follows. When $U' \equiv 0$, the waves have a forward motion under the density interface at the wave crest (phase zero) and a backward motion above and this flow is reversed at the wave trough. The net flow is zero and so, since the greatest forward flow is below the interface at phase zero, it is therefore here that the particle speed first equals the phase speed as the slope is increased. A small increase in δ reduces the forward flow below the interface and a wave of greater amplitude is needed to achieve conditions for wave breaking. The reverse is true for small negative δ. The abrupt decrease in critical slope of the first-order curve near $|\delta| = 0{\cdot}7$ in figure 7 is associated with the approach of c to U_0 (as in figure 1), and the redistribution of ψ (and therefore $\partial\psi/\partial z$, the horizontal current) and η as δ changes. The structure of ψ (see figure 2) becomes complex with regions of alternating horizontal currents at large $|\delta|$.

At second order the abrupt changes in phase and density level at which the phase speed equals $\max u$ can be seen to be directly related but in a way which it is difficult to interpret physically. At third order a more coherent pattern has emerged with the phase changing from π to 0 near $\delta = 0$ and a general reduction in level (figure 8*a*) except near $\delta = 0$, where the physical mechanism which dominated the first-order solution there continues to influence the solution.

Figure 1 shows that $c - U_0$ decreases as $|\delta|$ increases. The difference between the first-order phase speed c and the mean flow speed U_0 when the breaking condition is satisfied at order ϵ^3 derives almost equally from the particle motion and the reduction in phase speed through the terms $\epsilon^3 c_2$, the latter dominating near $|\delta| = 0{\cdot}6$, where c_2/c has its largest negative values. The nonlinearity of the system thus plays an inherently important part in providing the conditions at which breaking occurs.

The Richardson number in the wave when the phase speed equals $\max u$ was calculated. Although Ri is invariably small in some part of the flow where the density gradient is weak, it remains greater than 0·25 at the level of largest density gradient, $z = 0{\cdot}25h$, until approximately $|\delta| = 0{\cdot}75$, i.e. just before the minimum Richardson number in the *mean* flow becomes equal to 0·25 at $|\delta| = 0{\cdot}78$. Infinitesimal waves of length equal to that of the internal wave become unstable in the shear flow when the Richardson number falls below 0·093, i.e. when $|\delta| > 1{\cdot}28$ (see Thorpe 1971, figure 2; the waves correspond to $\alpha = \frac{1}{30}\pi$).

6. Experiments on breaking waves

The tube was filled as before with brine and water with $\alpha = \frac{1}{4}h$, $\Delta = (25{\cdot}0 \pm 0{\cdot}4) \times 10^{-3}$ and $\beta h = 32{\cdot}2 \pm 1{\cdot}46$, close to the value chosen for numerical study. The experiments were made with waves of the first mode having a variety of wave amplitudes, from the smallest which could be obtained to the largest for which no mixing was produced by the wave maker. The wave frequency was $1{\cdot}85 \pm 0{\cdot}02$ rad s^{-1} and this gave a measured wavenumber k such that $kh = 3{\cdot}246 \pm 0{\cdot}103$. Once the waves were established, the tube was tilted from the horizontal by $(5{\cdot}08 \pm 0{\cdot}05) \times 10^{-2}$ rad and the evolution of the waves recorded by the cameras. The value $|\delta| = 0{\cdot}5$ is attained by the mean flow after about 1·08 wave periods. This was a shorter time than might have been ideal, but was a compromise to avoid significant viscous effects at the walls and interface (see § 4) and the premature interruption of the flow by the arrival of surges from the tube ends.

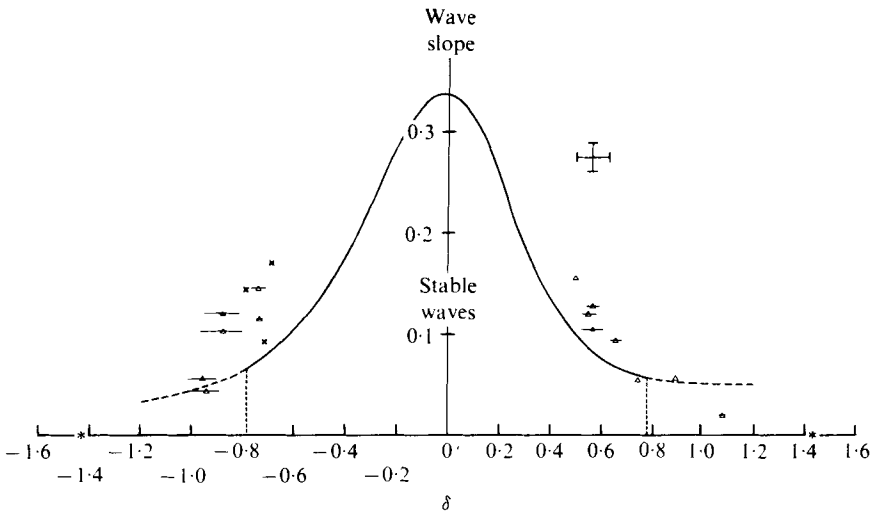


FIGURE 13. The variation of the wave slope (height × π/wavelength) with δ at the onset of wave breaking. The curve is the third-order curve of figure 7 and the points are from the experiments: Δ, average value determined from ciné and 'still' photographs with the experimental scatter; ×, values determined from ciné film only; *, value at the time when Kelvin–Helmholtz instability was first noticeable. The vertical cross gives estimated error bars for the experimental points.

Figures 9–12 (plates 2–9) show the appearance of the flow when wave breaking occurs. In the absence of waves Kelvin–Helmholtz instability, of course, is found (Thorpe 1971) and is first apparent at about $|\delta| = 1{\cdot}43$. When large internal waves were present, however, it was observed that they became asymmetrical and sharp at their crests (for positive δ, figures 9 and 11) or troughs (for negative δ, figures 10 and 12) where the vertical density gradients were reduced, and that this part of the wave overtook the sloping region in front of it in a manner similar to the commencement of breaking as surface waves approach a gradually shelving beach. Instead of plunging rapidly however (or, for negative δ, of rising), the sheared-off fluid moved forward with the flow at its own level, forming a horizontal plate with a well-defined nose. This development can be seen in the figures. (The breaking at negative δ in figure 10 may be more easily seen if the figure is inverted.) The remainder of the wave surface for positive δ became flattened and the wave decreased rapidly in amplitude, leaving the appearance of a layered structure. This, and the shock-like nose of the plate suggest a (catastrophic) transition from a dispersive wave structure to a non-linear shock or hyperbolic regime. The later development of the plate was interrupted by the onset of Kelvin–Helmholtz instability. For negative δ, the wave breaking was delayed (it occurred at larger $|\delta|$, see figure 13) and the plate motion was interrupted at an earlier stage by Kelvin–Helmholtz instability. The latter was usually modulated, the largest and first waves growing in the region of the plate, although the time at which the instability was first detectable was not significantly different from the value $|\delta| = 1{\cdot}43$ observed in the absence of waves. The wavelength of the growing billows accompanying instability was slightly modulated by the internal waves, but in no case was the billow wavelength equal to the wavelength of the free internal waves. The plates of fluid displaced from the broken internal waves were drawn and entrained into the spirals of the growing billows.

The early stages of the internal wave breaking, when asymmetry was apparent, were accompanied by the appearance of thin (2 mm vertical scale) laminae near the crest (or trough), possibly the result of the parametric instability described by McEwan & Robinson (1975). At very low wave slopes ($< 0{\cdot}05$) a rapid increase in the intensity of the thin layers seen on the shadowgraph preceded Kelvin–Helmholtz instability and was taken as evidence of the onset of wave breaking, but no region of positive vertical density gradients was apparent.

It is in practice difficult to determine with any degree of precision the exact moment at which waves begin to break. Breaking was examined both in the sequence of 'still' and in the sequence of ciné photographs to determine the time of its first occurrence and thus a measure of U_0 [equation (24)] and δ, and some effort was made to reduce subjective influence by averaging separate estimates. The wave slope was estimated simply from the crest-to-trough height and wavelength measured from the films, and no attempt was made to correct for the third-order contribution to the amplitude neglected in figure 7. The results are shown in figure 13 compared with the third-order theoretical curve from figure 7. The points lie almost entirely in the unstable region above the curve as expected, and their trend roughly follows that of the curve, reproducing quite well the asymmetry for positive and negative δ. For $\delta < 0$ breaking occurred in every case at or near the wave trough (phase π) whilst for $\delta > 0$ the breaking was first observed near the wave crest (phase zero), in agreement with the third-order prediction of figure 8 (b) for the range of slopes attained in the experiment. The fluid involved in the breaking appears to originate from above (for $\delta > 0$) or below ($\delta < 0$) the main density interface in accordance with figure 8 (a). The weakness of vertical gradients of density in these regions explains the absence of a marked tendency for the fluid plate to recover its own density level z_0. At the smallest slopes it was particularly difficult to estimate both the slope and the onset of breaking, and it is uncertain whether the points lying below the curve indicate any lack in the theoretical prediction.

7. Discussion

We have examined the effect of shear on finite amplitude internal waves, developing a theory which is valid for waves of small, but finite, amplitude and limited wavelength (these limitations are discussed in I), and deriving numerically some results in a particular case which it was possible to simulate approximately in an experiment. The observations on the shape of the waves supplement the conclusions of Long (1972), who examined long waves in a shear flow. A more detailed study of the way in which the second-harmonic contribution to the shape changes its phase in the accelerating flows is needed.

Perhaps the most interesting aspect is that of wave breaking. This clearly needs more attention. The laboratory experiments reported here go little further than an examination of the onset of breaking and more are needed, particularly in fluids with different density and velocity distributions, to ascertain the development of the overturned region, the possibility of rotor formation, and the energetics of the process. It is remarkable that when breaking occurs, at least at positive δ (see, for example, figure 9), the wave amplitude is so rapidly reduced. The largest particle displacements become concentrated near the breaking level (see figure 2) and presumably (although this is not clear from the ψ profiles in figure 2) the kinetic energy of the wave is also

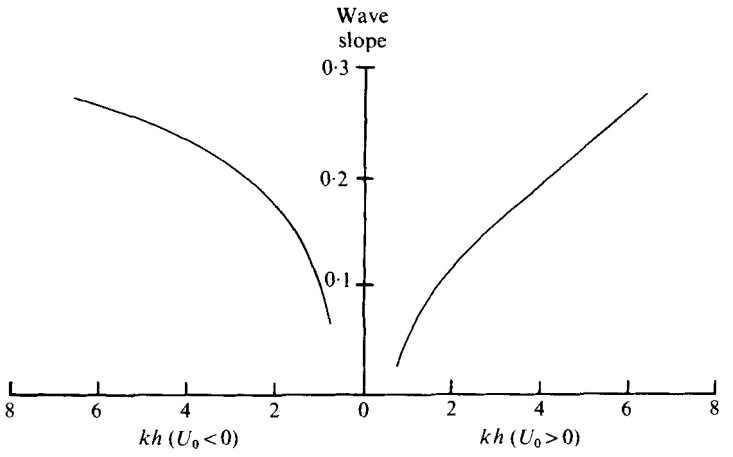


FIGURE 14. The variation with the wavenumber k of the smallest slope at which wave breaking can occur in waves in a fluid which is given by (23) with $\beta h = 20$, $d = 0{\cdot}25h$ and has a minimum Richardson number of 1·12.

a maximum and is concentrated near this level, so that when breaking occurs the plate of fluid which breaks away from the wave crest carries with it much of the available energy contained in the wave, both potential and kinetic. How this is eventually dissipated or transferred to mean flow and mean potential energy is not yet clear. The effects of the Reynolds number of the flow and the Rayleigh number in the overturning region of the wave need consideration if the results are to be applied. It is however clear that the presence of shear may very significantly reduce the slope (and wave energy) necessary for wave breaking. The mechanism of breaking may provide a clue to momentum and heat transfer across a turbulent–laminar interface at large Richardson numbers, as studied for example by Kantha, Phillips & Azad (1977). In geophysical applications the shear might be provided by a low-frequency wave component, and this work may thus be regarded as part of a more general study of wave pair interactions.

The effect of the wavelength may also be important in determining the slope at which waves will first break. Figure 14 shows the slope, calculated at third order, at which waves of the first mode in the tanh velocity and density profiles break, when the minimum mean flow Richardson number is kept constant at 1·12 and the wavelength is varied. The figure shows that the longer waves break at smaller slopes than the short waves, even though the former are moving more rapidly and might therefore be expected to be less affected by the mean shear, which is kept constant. The energy density of the waves required to achieve breaking is however greater the longer the wave, but bears no simple power-law relation to the wavenumber.

In the laboratory experiments wave breaking precedes Kelvin–Helmholtz instability. Whilst this is not true in general it suggests that a careful assessment needs to be made of the conditions under which shear-flow instability will occur, particularly in the ocean and elsewhere when, as is demonstrated here, the minimum Richardson number in the flow is not found at the depth where the density gradient is greatest. Miropolskiy & Filyushkin (1971) found that the observed slopes of fairly high frequency internal waves in the ocean were less than those needed to produce a self-induced

shear instability, so that Phillips' (1966, p. 188) model spectrum for 'saturated' waves was not appropriate. Frankignoul (1972) showed that wave breaking should precede Kelvin–Helmholtz instability in his simple wave shear model (results which are in accordance with those reported here), except for waves near the inertial frequency, and Garrett & Munk (1972) showed that indeed Kelvin–Helmholtz instability is more likely in the ocean since the energy of the waves is concentrated near inertial frequencies. The present results emphasize the importance of nonlinear effects and mean shear, particularly in the seasonal thermocline, where both effects may be most significant. The relative importance of wave breaking and Kelvin–Helmholtz instability, especially as diffusion processes, is yet uncertain.

The effect of shear on the breaking of internal waves is exactly analogous to that of wind drift on surface waves studied by Phillips & Banner (1974), resulting in an increase in particle speeds and a tendency to initiate breaking. The wind-driven mixing layer at the top of the ocean may be regarded as a potential site of internal wave breaking, particularly when the layer is accelerated by momentum transferred to it from the wind. The resulting breaking, being asymmetrical according to the direction of propagation of the waves relative to the mean shear, will modify an existing field of internal waves, and possibly leave a wave field with directional asymmetry. It is doubtful whether such effects could in practice be distinguished in periods of active atmospheric forcing, because of the variety of other mechanisms which might then modify the wave spectrum, but the physical features accompanying wave breaking might be observed and experiments are planned to investigate the detailed structure of the mixing layer and upper thermocline.

I am grateful to Alan Hall and Mike Bray for their help in making the laboratory experiments.

Appendix A. The two-layer fluid

Consider two layers of fluid, the upper of density ρ_1, depth h_1 and moving at speed U_1, and the lower of density ρ_2, depth h_2 and moving at speed U_2 in the same direction, the two layers being confined between horizontal boundaries at $z = -h_2$, h_1. It may easily be shown by a conventional Stokes-type expansion that the displacement of the interface by a two-dimensional wave of wavenumber k moving with speed c parallel to the motion in the layers (which are supposed inviscid and irrotational) is

$$\eta = A\cos k(x-ct) + \frac{A^2k}{4t_1^2t_2^2}\left[\frac{\rho_2 c_2^2 t_1^2(3-t_2^2)-\rho_1 c_1^2 t_2^2(3-t_1^2)}{c_1^2\rho_1 t_1 + c_2^2\rho_2 t_2}\right]\cos 2k(x-ct)$$

plus higher-order terms in the 'wave slope' Ak, where

$$c_1^2\rho_1 t_2 + c_2^2\rho_2 t_1 = (\rho_2-\rho_1)gt_1t_2/k,$$

and $t_i = \tanh kh_i$, $c_i = c - U_i$ $(i = 1, 2)$. The dispersion relation may, for certain values of k, U_1 and U_2, lead to imaginary values of c, and thus unstable, growing waves.

The distortion of the wave shape η from sinusoidal, given by the coefficient of $\cos 2k(x-ct)$, may be positive (as for surface waves) or negative. If, for example (as

in the main text), $k = \pi/(h_1+h_2)$, $h_1 = 3h_2$, $U_i = -(-1)^i U_0$ and $\rho_2 - \rho_1 \ll \rho_1$, then the coefficient is positive for $U_* = -0{\cdot}26c_0 < U_0 < 1{\cdot}02c_0$ but negative for

$$-1{\cdot}02c_0 < U_0 < U_*,$$

where c_0 is the speed of waves of wavenumber k in the absence of shear, i.e. when $U_0 = 0$. For U_0 outside this range, waves of wavenumber k are unstable.

Appendix B. Expressions for $\sigma_2, \phi_{33}, \theta_1, \theta_2, \theta_3, \theta_4$

If $s = U - c$ and $r = g\rho_0'/\rho_0(0)$,

$$\sigma_2 = \frac{(c-c_g)}{k\int_0^h \Psi^2\,dz}\left[\int_0^h \frac{\Psi^2\Psi_2}{s}\left(\frac{3rs'}{s^3} - \frac{2r'}{s^2} - \left(\frac{s''}{s}\right)'\right)dz\right.$$

$$\left. - \int_0^h \Psi^2\Psi'^2\frac{r}{s^4}\,dz + \int_0^h \frac{\Psi^4}{4s^6}(3rs'^2 - 4rs^2k^2 - 4r^2 - (r+s'^2)ss'' + s^2s's'')\,dz\right]$$

and

$$\phi_{33} = A^3\Psi_3(z),$$

where

$$\mathscr{L}(3k,c)\Psi_3 = \frac{\Psi^3}{2s^3}\left[\frac{s^{(\mathrm{iv})}}{3} + \frac{1}{s}\left(r'' - s's''' - \frac{s''^2}{3}\right) + \frac{1}{s^2}\left(s'^2s'' - \frac{4}{3}rs'' - 4r's'\right) + \frac{5rs'^2}{s^3}\right] + \frac{4\Psi\Psi_2}{s^2}\left[\frac{s'''}{2} + \frac{1}{s}\left(r' - \frac{s's''}{2}\right) - \frac{4rs'}{3s^2}\right].$$

The functions appearing in the particle speed are

$$\theta_1(z) = \frac{\Psi_0'''\psi_1^2}{4s^2} + \frac{\psi_1'^2 - \psi_1\psi_1''}{2s} + \psi_2' + \frac{\Psi_0''}{2s}\left(\frac{\psi_1^2\Psi_0''}{2s^2} - \frac{\psi_1\psi_1'}{s} - 2\psi_2\right),$$

$$\theta_2(z) = \frac{\Psi_0'''\psi_1^2}{4s^2} - \frac{\psi_1'^2 + \psi_1\psi_1''}{2s} + \frac{\Psi_0''\psi_1\psi_1'}{2s^2},$$

$$\theta_3(z) = \frac{\Psi_0''}{s}\left\{\frac{3}{8}\psi_1 x_{11}^2 + 2\psi_2 x_{11} - \frac{x_{21}\psi_1}{2} - \frac{z_{11}\psi_1'}{4}x_{11} - \frac{z_{11}^2\psi_1''}{8} - \psi_2' z_{11}\right\} + \frac{z_{11}z_{21}}{2}\Psi_0''' + \frac{z_{11}^3}{8}\Psi_0^{(\mathrm{iv})} + z_{21}\frac{\psi_1''}{2} + \frac{3}{8}z_{11}^2\psi_1''' - \frac{x_{11}z_{11}\psi_1''}{4} - \frac{\psi_1' x_{21}}{2} - \frac{x_{11}^2\psi_1'}{2} - x_{11}\psi_2' + \frac{z_{11}\psi_2''}{2}$$

and

$$\theta_4(z) = -\frac{\Psi_0''}{3s}\left\{\frac{\psi_1 x_{21}}{2} + \frac{\psi_1 x_{11}^2}{8} + \frac{z_{11}x_{11}\psi_1'}{4} + \frac{z_{11}^2\psi_1''}{8} + 2\psi_2 x_{11} + \psi_2' z_{11} + 3\psi_3\right\} + \frac{z_{11}z_{21}}{2}\Psi_0''' + \frac{z_{11}^3\Psi_0^{(\mathrm{iv})}}{24} + \frac{z_{21}\psi_1''}{2} + \frac{z_{11}^2\psi_1'''}{8} + \frac{x_{11}z_{11}\psi_1''}{4} + \frac{\psi_1' x_{21}}{2} + \frac{x_{11}^2\psi_1'}{2} - x_{11}\psi_2' + \frac{z_{11}\psi_2''}{2} + \psi_3',$$

where

$$x_{11} = \frac{1}{s}\left(\psi_1' - \frac{\psi_1\Psi_0''}{s}\right), \quad z_{11} = -\frac{\psi_1}{s},$$

$$x_{21} = \frac{\theta_1}{2s}, \quad z_{21} = \frac{1}{2s}\left(\frac{\psi_1^2\Psi_0''}{2s^2} - 2\psi_2\right).$$

REFERENCES

BANKS, W. H. H., DRAZIN, P. G. & ZATURSKA, M. B. 1976 On the normal modes of parallel flow of inviscid stratified fluid. *J. Fluid Mech.* **75**, 149–171.

BENJAMIN, T. B. 1966 Internal waves of finite amplitude and permanent form. *J. Fluid Mech.* **25**, 241–270.

BORISENKO, YU. D., VORONOVICH, A. G., LEONOV, A. I. & MIROPOLSKIY, YU. Z. 1976 Towards a theory of non-stationary weakly non-linear internal waves in a stratified fluid. *Izv. Atoms. Ocean Phys.* **12**, 174–179 (English edn).

CHANDRASEKHAR, S. 1961 *Hydrodynamic and Hydromagnetic Stability.* Oxford: Clarendon Press.

DAVEY, A. & STEWARTSON, K. 1974 On three-dimensional packets of surface waves. *Proc. Roy. Soc.* A **338**, 101–110.

ELIASSEN, A., HØILAND, E. & RIIS, E. 1953 Two-dimensional perturbation of a flow with a constant shear of a stratified fluid. *Inst. Weather Climate Res., Oslo, Publ.* no. 1.

FRANKIGNOUL, C. J. 1972 Stability of finite amplitude internal waves in a shear flow. *Geophys. Fluid Dyn.* **4**, 91–99.

GARRETT, C. & MUNK, W. 1972 Oceanic mixing by breaking internal waves. *Deep-Sea Res.* **19**, 823–832.

GRIMSHAW, R. 1976 The modulation and stability of an internal gravity wave. *Mem. Soc. Roy. Sci. Liege, Ser.* 6, **10**, 299–314.

HAZEL, P. 1972 Numerical studies of the stability of inviscid stratified shear flows. *J. Fluid Mech.* **51**, 39–62.

HOWARD, L. N. 1963 Neutral curves and stability boundaries in stratified flow. *J. Fluid Mech.* **16**, 333–342.

KANTHA, L. H., PHILLIPS, O. M. & AZAD, R. S. 1977 On turbulent entrainment at a stable density interface. *J. Fluid Mech.* **79**, 753–768.

KEULEGAN, G. H. & CARPENTER, L. H. 1961 An experimental study of internal progressive oscillatory waves. *Nat. Bur. Stand. Rep.* no. 7319.

LONG, R. R. 1972 The steepening of long internal waves. *Tellus* **24**, 88–99.

MCEWAN, A. D. & ROBINSON, R. M. 1975 Parametric instability of internal gravity waves. *J. Fluid Mech.* **67**, 667–688.

MILES, J. W. 1967 Internal waves in a continuously stratified atmosphere or ocean. *J. Fluid Mech.* **28**, 305–310.

MIROPOLSKIY, YU. Z. & FILYUSHKIN, B. N. 1971 Temperature fluctuations in the upper ocean comparable in scale to internal gravity waves. *Izv. Atmos. Ocean Phys.* **7**, 523–535 (English edn).

ORLANSKI, I. & BRYAN, K. 1969 Formation of the thermocline step structure by large amplitude internal gravity waves. *J. Geophys. Res.* **74**, 6975–6983.

PHILLIPS, O. M. 1966 *The Dynamics of the Upper Ocean.* Cambridge University Press.

PHILLIPS, O. M. & BANNER, M. L. 1974 Wave breaking in the presence of wind drift and swell. *J. Fluid Mech.* **66**, 625–640.

STOKES, G. G. 1947 On the theory of oscillatory waves. *Trans. Camb. Phil. Soc.* **8**, 441–455.

STUART, J. T. 1960 On the non-linear mechanics of wave disturbances in stable and unstable parallel flows. Part 1. *J. Fluid Mech.* **9**, 353–370.

TAYLOR, G. I. 1931 Effect of variation of density on the stability of superposed streams of fluid. *Proc. Roy. Soc.* A **132**, 499–523.

THORPE, S. A. 1968*a* On the shape of progressive internal waves. *Phil. Trans. Roy. Soc.* A **263**, 563–614.

THORPE, S. A. 1968*b* A method of producing a shear flow in a stratified fluid. *J. Fluid Mech.* **32**, 693–704.

THORPE, S. A. 1969 Neutral eigensolutions of the stability equation for stratified shear flow. *J. Fluid Mech.* **36**, 673–683.

THORPE, S. A. 1971 Experiments on the instability of stratified shear flows: miscible fluids. *J. Fluid Mech.* **46**, 299–319.

THORPE, S. A. 1974 Near-resonant forcing in a shallow two-layer fluid: a model for the internal surge in Loch Ness? *J. Fluid Mech.* **63**, 509–527.

THORPE, S. A. 1977 On the stability of internal wavetrains. In *Deep-Sea Res., A Voyage of Discovery* (ed. M. V. Angel), pp. 199–212. Pergamon.

VORONOVICH, A. G. 1976 Propagation of internal and surface gravity waves in the approximation of geometrical optics. *Izv. Atmos. Ocean Phys.* **12**, 519–523 (English edn).

YIH, C. S. 1974 Progressive waves of permanent form in continuously stratified fluids. *Phys. Fluids* **17**, 1489–1495.

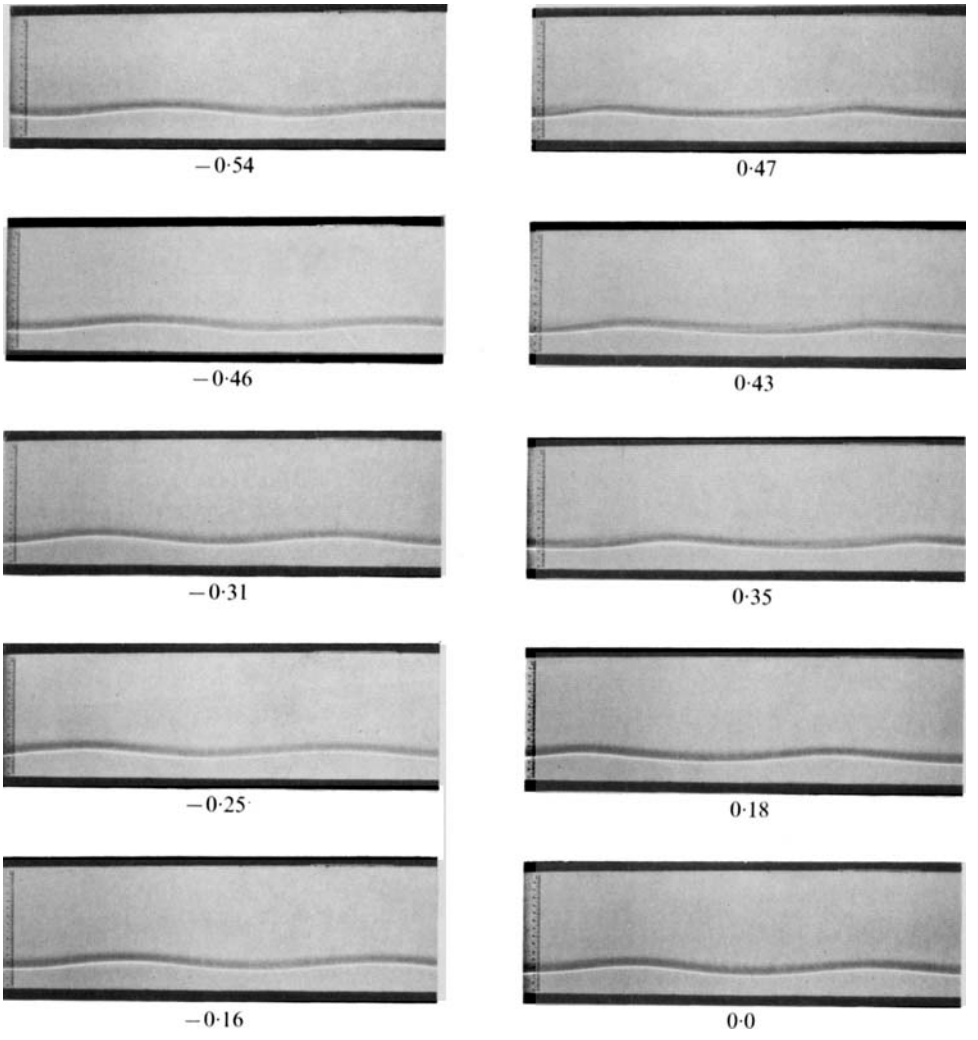
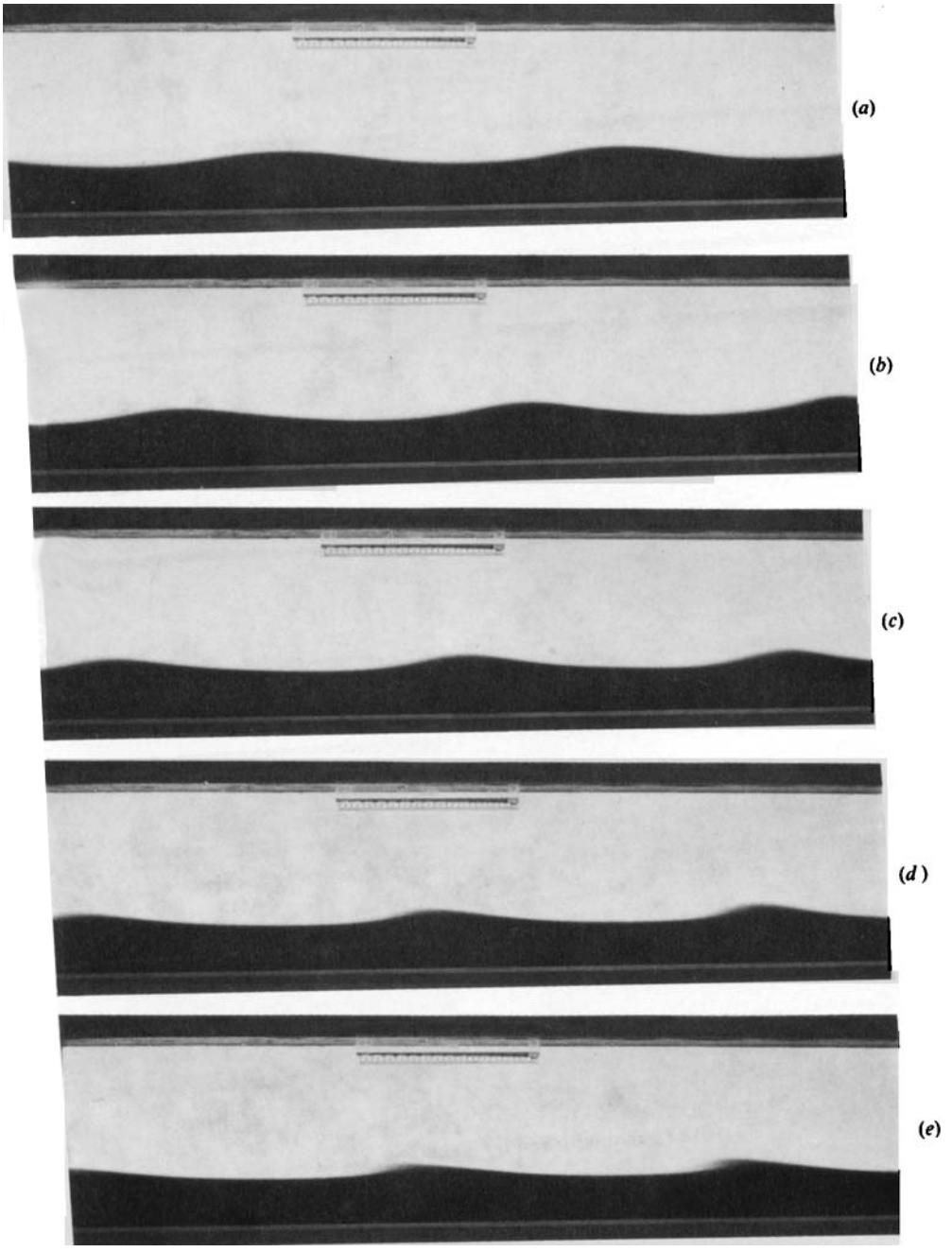


FIGURE 6. The shape of internal waves at various values of δ made visible by a shadowgraph.



FIGURES 9 (*a–e*). For legend see facing page.

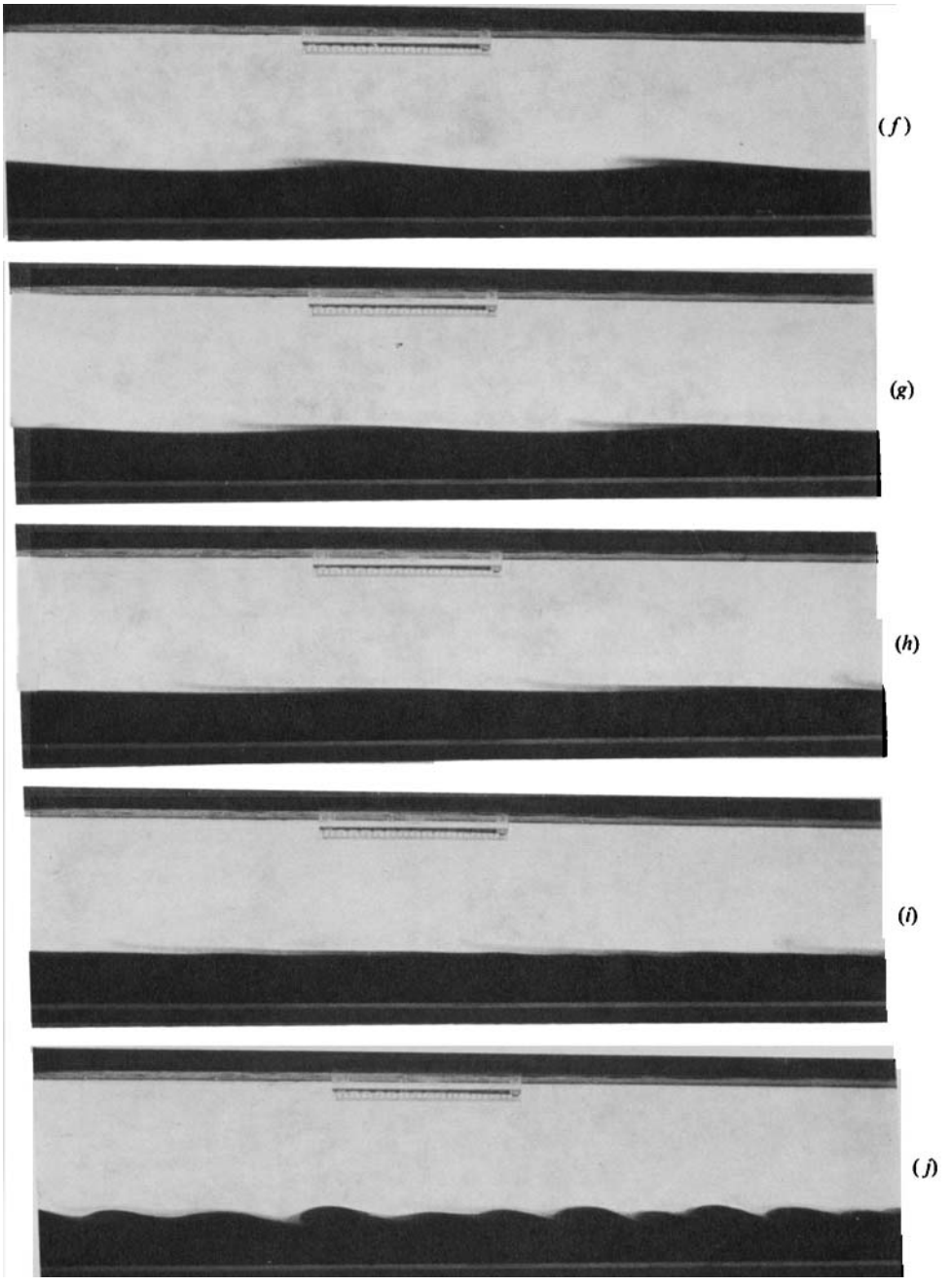
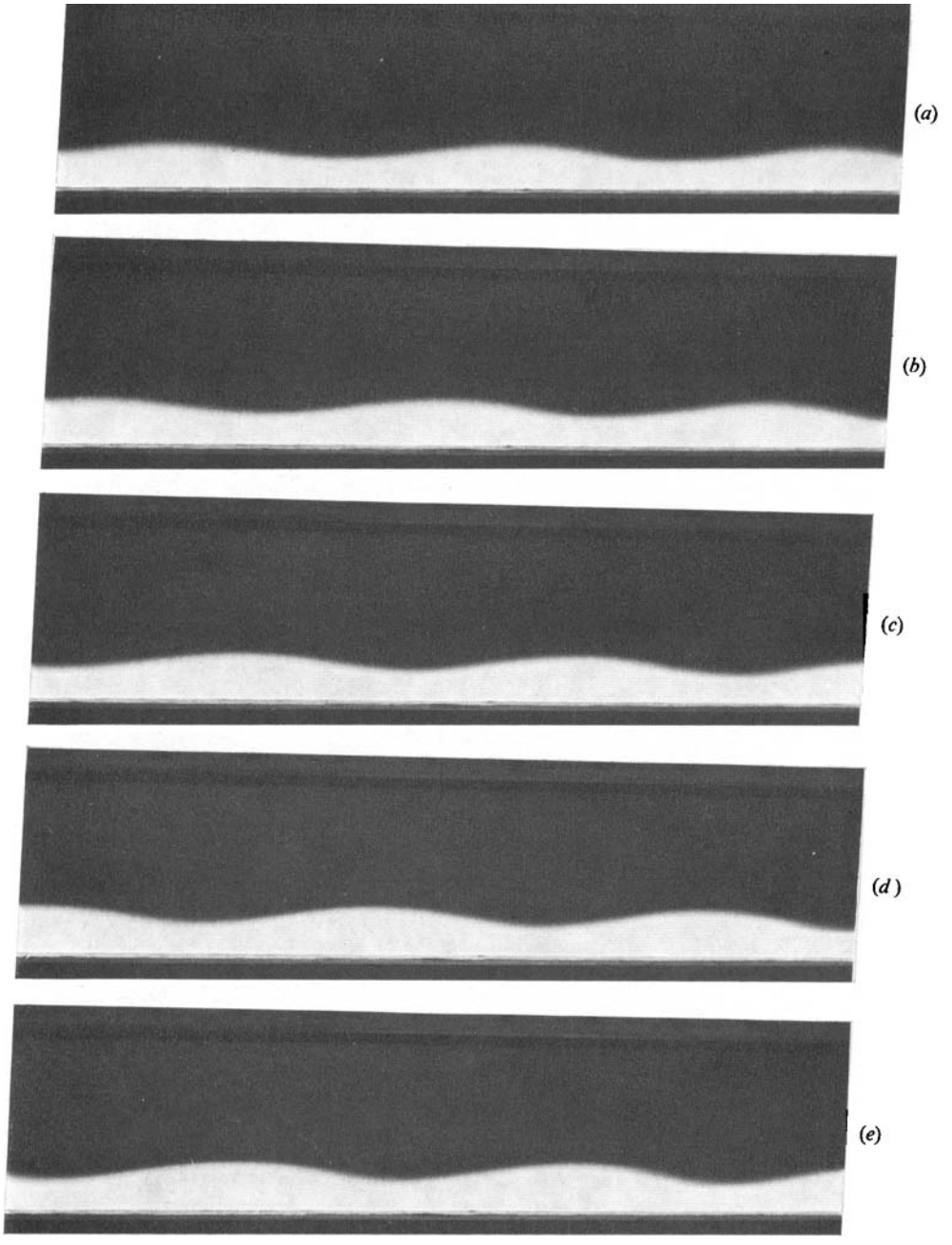


FIGURE 9. The onset of breaking for positive δ. The waves are moving to the left. The lower brine layer has been dyed with potassium permanganate. The photos were taken at the following times after the right end of the tube was tilted down: (*a*) 1·8 s; (*b*) 3·3 s; (*c*) 4·7 s; (*d*) 5·6 s; (*e*) 6·6 s; (*f*) 7·6 s; (*g*) 8·5 s; (*h*) 9·5 s; (*i*) 10·5 s; (*j*) 11·4 s. The onset of Kelvin–Helmholtz instability is seen in (*j*).



FIGURES 10 (*a–e*). For legend see facing page.

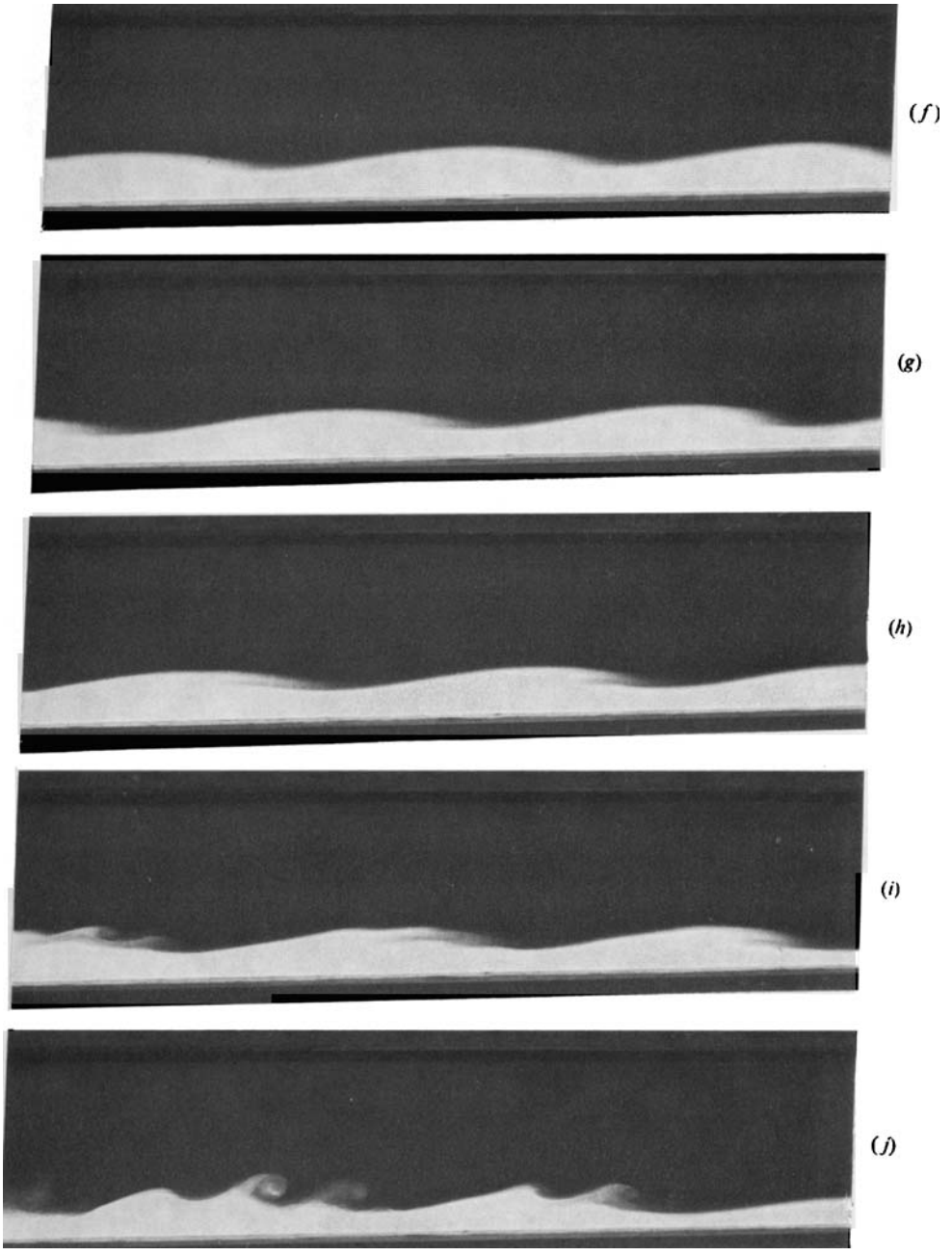
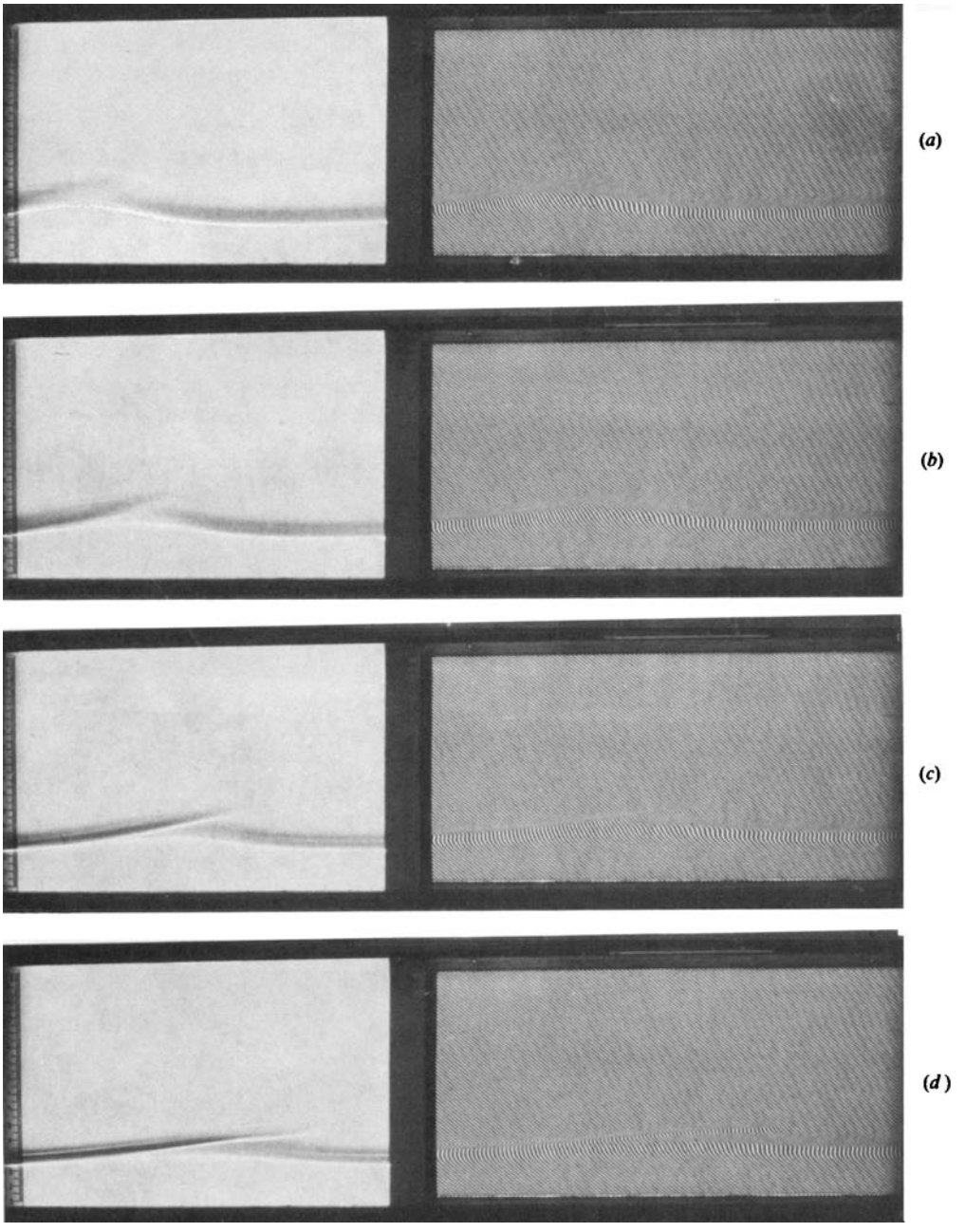


FIGURE 10. The onset of breaking for negative δ. The waves are moving to the left. The upper brine layer has been dyed with potassium permanganate. The photos were taken at the following times after the left end of the tube was tilted down: (*a*) 2·0 s; (*b*) 3·4 s; (*c*) 4·8 s; (*d*) 6·2 s; (*e*) 7·2 s; (*f*) 8·2 s; (*g*) 9·1 s; (*h*) 10·1 s; (*i*) 11·0 s; (*j*) 12·0 s. Kelvin–Helmholtz instability is seen in (*i*) and (*j*).



FIGURES 11 (*a–d*). For legend see facing page.

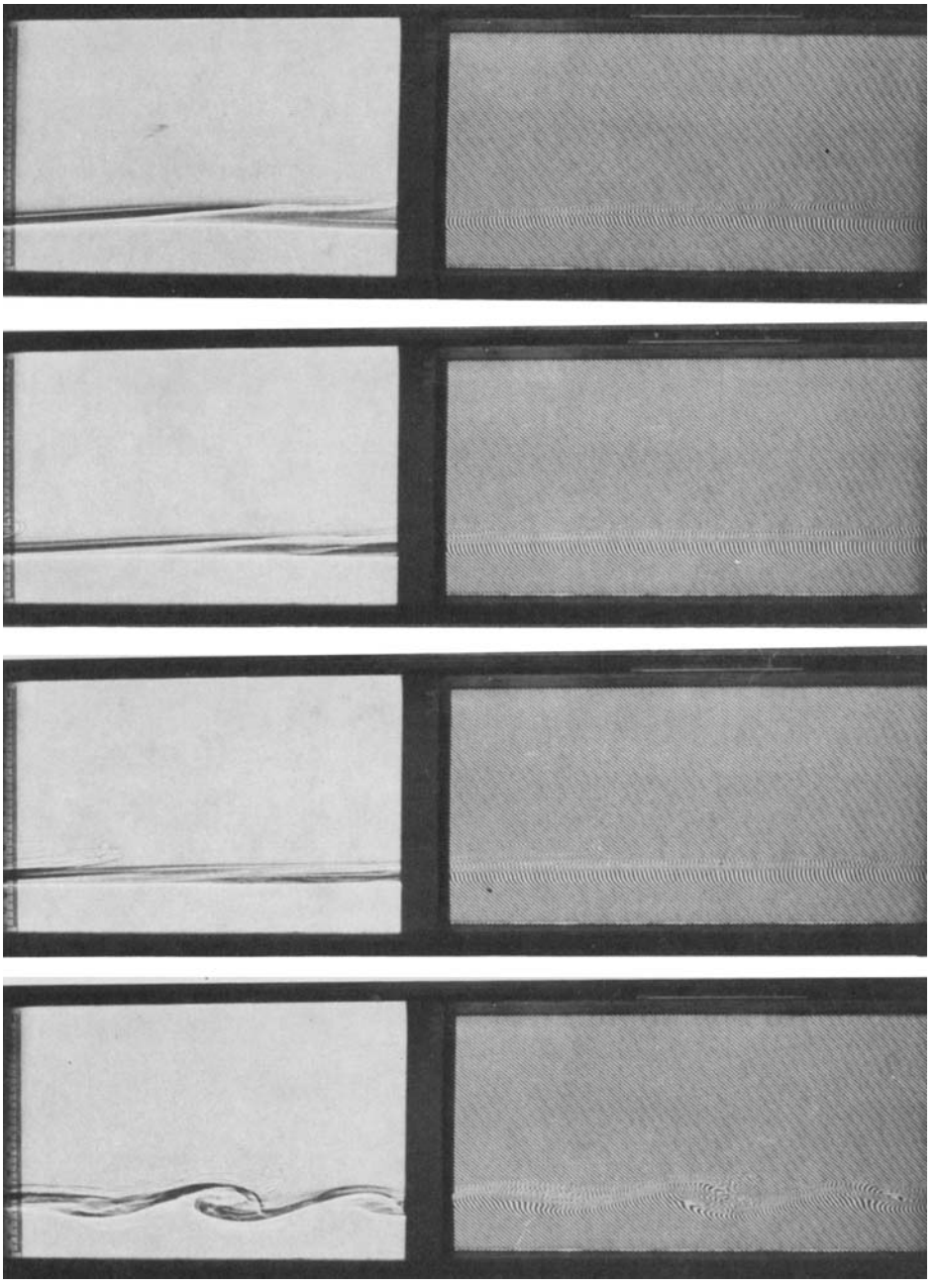
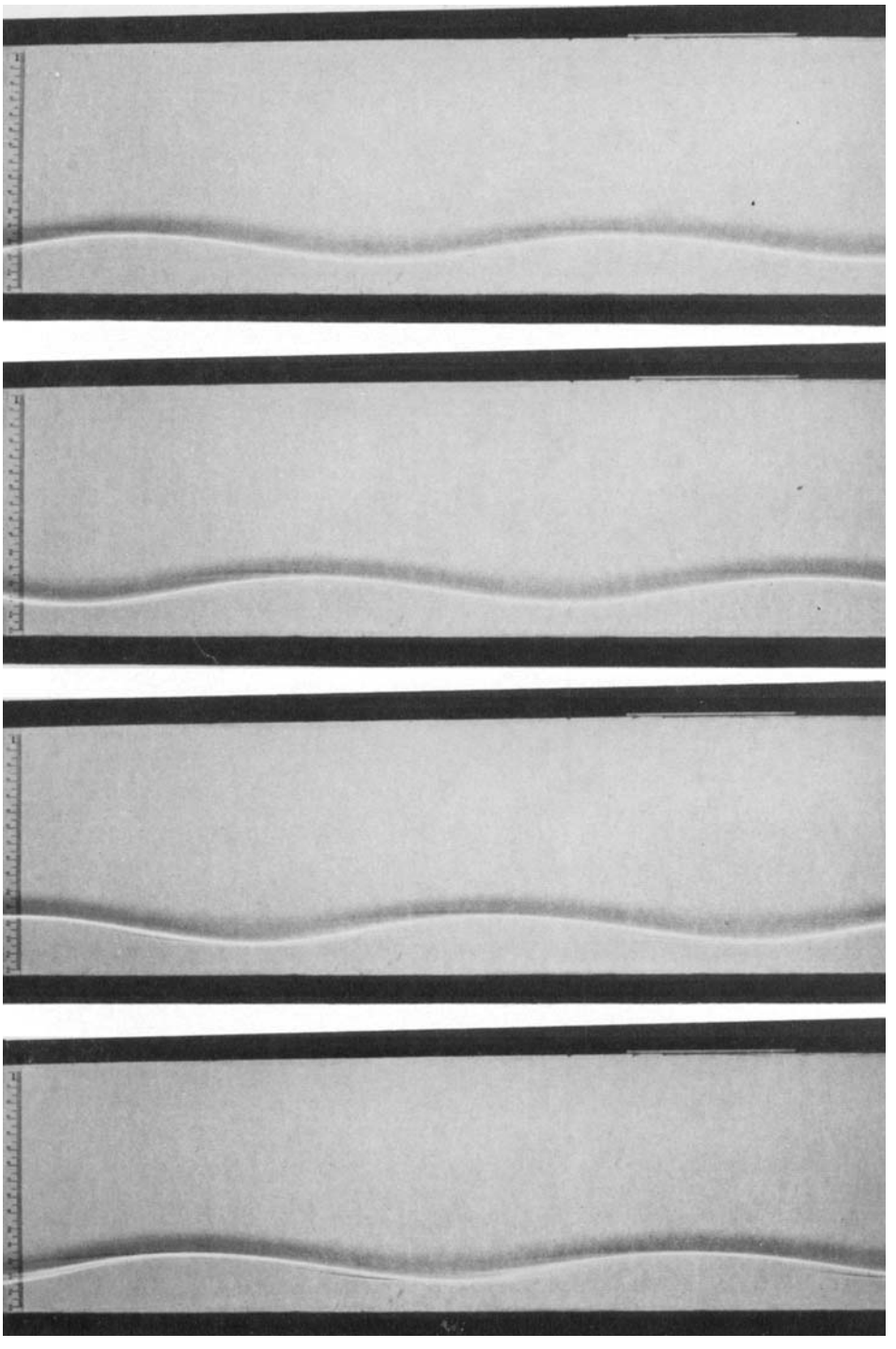


FIGURE 11. The onset of breaking for positive δ. The waves are moving to the right, and are made visible at the left by a shadowgraph and on the right by the distortions they make in a set of parallel lines viewed through the tube. The photos were taken at the following times after the left end of the tube was tilted down: (*a*) 5·2 s; (*b*) 6·2 s; (*c*) 7·1 s; (*d*) 8·0 s; (*e*) 9·0 s; (*f*) 9·9 s; (*g*) 10·9 s; (*h*) 12·6 s.



FIGURES 12 (*a–d*). For legend see facing page.

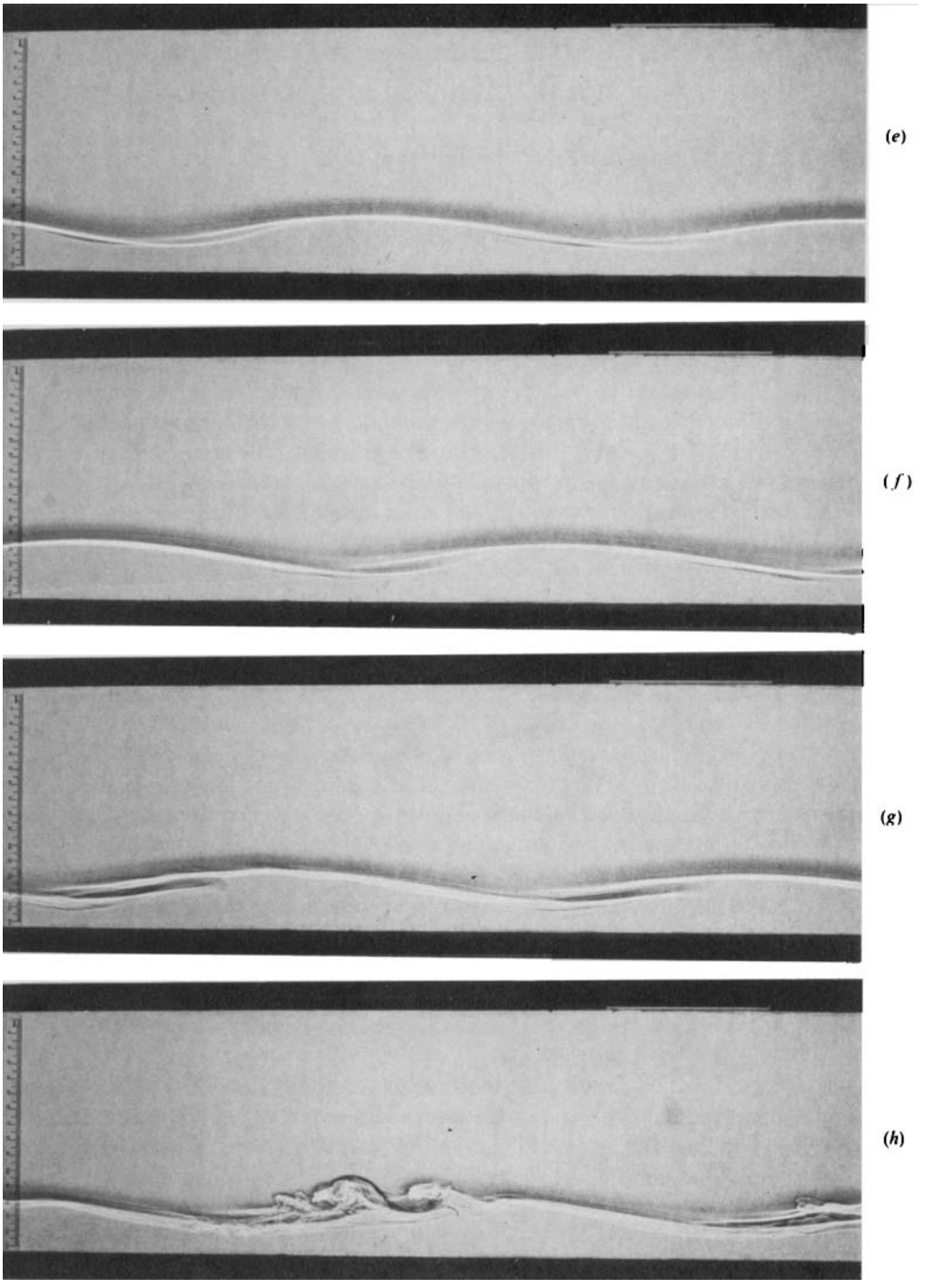


FIGURE 12. The onset of breaking for negative δ. The waves are moving to the right and are made visible by a shadowgraph. The photos were taken at the following times after the right end of the tube was tilted down: (*a*) 3·6 s; (*b*) 4·6 s; (*c*) 5·6 s; (*d*) 6·5 s; (*e*) 7·5 s; (*f*) 8·4 s; (*g*) 9·4 s; (*h*) 10·3 s.