J. Fluid Mech. (1967), *vol.* 27, *part* 2, *pp.* 399–412
Printed in Great Britain

Non-linear dispersion of water waves

By G. B. WHITHAM
California Institute of Technology

(Received 25 February 1966)

The slow dispersion of non-linear water waves is studied by the general theory developed in an earlier paper (Whitham 1965*b*). The average Lagrangian is calculated from the Stokes expansion for periodic wave trains in water of arbitrary depth. This Lagrangian can be used for the various applications described in the above reference. In this paper, the crucial question of the 'type' of the differential equations for the wave-train parameters (local amplitude, wave-number, etc.) is established. The equations are hyperbolic or elliptic according to whether κh_0 is less than or greater than 1·36, where κ is the wave-number per 2π and h_0 is the undisturbed depth. In the hyperbolic case, changes in the wave train propagate and the characteristic velocities give generalizations of the linear group velocity. In the elliptic case, modulations in the wave train grow exponentially and a periodic wave train will be unstable in this sense; thus, periodic wave trains on water will be unstable if $\kappa h_0 > 1{\cdot}36$. The instability of deep-water waves, $\kappa h_0 \to \infty$, was discovered in a different way by Benjamin (1966). The relation between the two approaches is explained.

1. Introduction

In recent papers (Whitham 1965*a*, *b*), a general theory has been given for the slow dispersion of non-linear wave trains. A uniform periodic wave train is specified by certain parameters such as amplitude, wave-number, etc.; the theory treats non-uniform wave trains in which these parameters vary slowly in space and time, in the sense that the changes in one wavelength and in one period are relatively small. All the problems considered to be typical in this theory stem from variational principles, and the mathematics is tremendously simplified by suitable use of the corresponding Lagrangians.

The Korteweg–de Vries and Boussinesq equations for long water waves were included as examples in the earlier papers, but it was not obvious that the full equations of water waves would fit into the same pattern. In the approximations for long waves, the dependence of the flow on the vertical co-ordinate y is eliminated to give equations for functions of the horizontal co-ordinate $\mathbf{x}$ and the time t. It was easy to spot the variational principles leading to these $(\mathbf{x}, t)$ equations. But, a corresponding variational principle for the full equations of water waves, including the free-surface conditions, did not seem to be known. It was also thought that the y-dependence, with the wave propagation occurring only in the $(\mathbf{x}, t)$ dependence, may require crucial changes in the general approach.

Now, Luke (1966) has shown that the equations of irrotational water waves follow completely from the variational principle

$$\delta \iint L\, d\mathbf{x}\, dt = 0,$$

$$L = \int_{-h_0}^{\eta} \{\phi_t + \tfrac{1}{2}(\nabla\phi)^2 + gy\}\, dy, \tag{1}$$

where $\phi(\mathbf{x}, y, t)$ is the velocity potential, $y = \eta(\mathbf{x}, t)$ is the free surface, $y = -h_0$ is the rigid bottom, and g is the acceleration of gravity. It turns out that this variational principle was given by Bateman (1944), but he did not note that the free-surface conditions (the essential difficulty for water waves) also follow. The general theory of dispersion can now be applied to this system with the Lagrangian given in (1) and follows the standard pattern. The y-dependence does not give any trouble; it is integrated out in (1).

Without the variational principle, one can take the required properties of uniform wave trains directly from the differential equations and then calculate a Lagrangian equal to the kinetic energy minus the potential energy, i.e.

$$L_1 = \int_{-h_0}^{\eta} \{\tfrac{1}{2}(\nabla\phi)^2 - gy\}\, dy, \tag{2}$$

and use this quantity to study the slowly varying wave trains. For any solution, (1) and (2) are equal, but varying (2) without further restrictions on ϕ yields only Laplace's equation and does not give the correct boundary conditions.

In this paper the equations for slow variations in the amplitude, wave-number etc. will be established for arbitrary depth. Of course, there is no explicit exact solution for the uniform wave train; it is calculated by the Stokes expansion in powers of amplitude. The theory for slowly varying wave trains will be worked out in the same way. The Stokes expansion breaks down for long waves, but these have been covered in the previous papers. The analysis is carried out for the case of one horizontal space dimension, but the extension to two dimensions is straightforward.

For finite depth, as in the long-wave approximation (Whitham 1965*b*), variations in mean height and mean fluid velocity occur and they are coupled nonlinearly with variations of the amplitude and wave-number. These four quantities are fundamental parameters of the wave train and ultimately a coupled set of four differential equations is obtained for them. The frequency and a sixth variable, a 'pseudo-frequency', are also fundamental in the analysis but are functionally related with the other four. Lighthill (1965), in his application of the theory, assumes from the outset that the mean height and mean velocity will play no role in the case of infinitely deep water. For long waves a similar assumption would lead to completely wrong results, so the present writer felt that a satisfactory treatment would have to wait for the complete analysis. It does turn out, however, that changes in mean height and velocity uncouple from the changes in amplitude and wave-number in the limit when the depth is large compared with the wavelength.

In the case of long waves, the equations for the slow variations are hyperbolic, showing that changes in the wave train propagate in a finite way. Lighthill's application of the theory to deep-water waves showed that the equations are elliptic in that case. This means, for example, that a small sinusoidal modulation of the amplitude or wave-number will grow in time, and in that sense the wave is unstable. This result was discovered in a different approach by Benjamin (1966) by a subtle use of the traditional approach to stability problems. His theory is limited to small modulations of a nearly linear main wave but it is not limited to slowly varying changes. Thus the modulation frequency does not need to be small. He finds a cut-off for the exponential growth when the difference between the two frequencies reaches a certain value.

The analysis given here for finite depth allows a study of the change of type of the dispersion equations from hyperbolic to elliptic equations as the ratio of depth to wavelength increases. The change occurs when $\kappa h_0 = 1{\cdot}36$, where κ is the wave-number per 2π, and h_0 is the undisturbed depth. The change of type depends crucially on the coupling between changes in mean depth and velocity with changes in wave-number and amplitude. The non-linear dependence of frequency on amplitude gives a contribution tending towards elliptic equations while the non-linear coupling gives a contribution in the hyperbolic direction.

2. The average Lagrangian for Stokes waves

The uniform periodic solution of the water wave equations takes the form

$$\left.\begin{aligned} \eta &= N(\theta), \quad \theta = \kappa x - \omega t, \\ \phi &= \beta x - \gamma t + \Phi(\theta, y), \end{aligned}\right\} \tag{3}$$

where κ, ω, β, γ are constant parameters. The pair (κ, ω) are the wave-number and frequency, and the phase function θ will be normalized so that it increases by 2π in one period. The linear term $(\beta x - \gamma t)$ must be allowed in ϕ, since it is only the derivatives of ϕ that represent periodic physical quantities. Physically β is the mean velocity, but the meaning of γ is less clear; it corresponds to absorbing the Bernoulli constant into the potential. Mathematically (β, γ) act like a pseudo wave-number and frequency in ϕ corresponding to the real wave-number and frequency (κ, ω) in θ. In the expressions for N and Φ there are two further main parameters taken to be the amplitude a and the mean value b of the height η. Thus, the solution depends upon two triads of parameters (κ, ω, a) and (β, γ, b) just as in the long-wave case (Whitham 1965*b*).

For non-uniform wave trains the form of the solution is generalized slightly to

$$\left.\begin{aligned} \eta &= N(\theta), \\ \phi &= \psi + \Phi(\theta, y), \end{aligned}\right\} \tag{4}$$

$$\left.\begin{aligned} \theta_x &= \kappa, \quad \theta_t = -\omega, \\ \psi_x &= \beta, \quad \psi_t = -\gamma, \end{aligned}\right\} \tag{5}$$

and the quantities (κ, ω, a), (β, γ, b) are all taken to be slowly varying functions of space and time instead of constant parameters. The equations for these functions are derived from an averaged form of the variational principle. The general method (Whitham 1965*b*) is to calculate the average Lagrangian over one wavelength, i.e.

$$\mathscr{L}(\kappa,\omega,a;\beta,\gamma,b) = \frac{1}{2\pi}\int_0^{2\pi} L\,d\theta, \tag{6}$$

from the uniform wave-train solution. Then, for the non-uniform wave train, where the scale is large compared with one wavelength, the averaged variational principle

$$\delta\iint \mathscr{L}\,dx\,dt = 0$$

is used, with $\kappa, \omega, \beta, \gamma$ restricted by (5). The Euler equations are

$$\mathscr{L}_a = 0, \quad \mathscr{L}_b = 0, \tag{7}$$

$$-\frac{\partial}{\partial t}\mathscr{L}_\omega + \frac{\partial}{\partial x}\mathscr{L}_\kappa = 0, \quad -\frac{\partial}{\partial t}\mathscr{L}_\gamma + \frac{\partial}{\partial x}\mathscr{L}_\beta = 0. \tag{8}$$

The system is completed by eliminating θ and ψ in (5) to give

$$\frac{\partial\kappa}{\partial t} + \frac{\partial\omega}{\partial x} = 0, \quad \frac{\partial\beta}{\partial t} + \frac{\partial\gamma}{\partial x} = 0. \tag{9}$$

The Stokes expansion for the uniform wave train is in powers of κa. The series is not uniformly valid as $\kappa h_0 \to 0$, however, and κa must be small compared with $(\kappa h_0)^3$ in that limit. The expansion is obtained from the Fourier series

$$N(\theta) = b + a\cos\theta + \sum_2^\infty a_n \cos n\theta, \tag{10}$$

and

$$\Phi = \sum_1^\infty \frac{A_n}{n}\cosh n\kappa(h_0+y)\sin n\theta, \tag{11}$$

where the expansion for Φ has been chosen to satisfy Laplace's equation and the boundary condition $\partial\Phi/\partial y = 0$ on the bottom $y = -h_0$. The mean height b and the amplitude a are considered to be the fundamental coefficients in this solution; the other coefficients a_n, A_n are determined in terms of them by satisfying the free-surface conditions. This can be done most conveniently, in the present context, by calculating the average Lagrangian (6) as a function of all the coefficients and using the variational equations

$$\mathscr{L}_{a_n} = 0, \quad \mathscr{L}_{A_n} = 0, \tag{12}$$

as well as (7) and (8). In the solution of these non-linear relations by successive approximations, it becomes clear that κa_n and κA_n are $O(\epsilon^n)$, where ϵ is a typical value of κa.

First, from (1) and (3),

$$L = (\tfrac{1}{2}\beta^2-\gamma)(h_0+N) + \tfrac{1}{2}gN^2 - (\omega-\beta\kappa)\int_{-h_0}^N \Phi_\theta\,dy + \int_{-h_0}^N (\tfrac{1}{2}\kappa^2\Phi_\theta^2 + \tfrac{1}{2}\Phi_y^2)\,dy. \tag{13}$$

After substituting the series (11) and carrying out the integrations with respect to y, we obtain

$$L = (\tfrac{1}{2}\beta^2-\gamma)(h+M)+\tfrac{1}{2}gb^2+gbM+\tfrac{1}{2}gM^2-\frac{\omega-\beta\kappa}{\kappa}\sum_{n=1}^{\infty}\frac{An}{n}\sinh n\kappa(h+M)\cos n\theta$$
$$+\tfrac{1}{4}\kappa\sum_{n=2}^{\infty}\sum_{r=1}^{n-1}A_rA_{n-r}\left\{\frac{\sinh n\kappa(h+M)}{n}\cos(n-2r)\theta\right.$$
$$\left.+\frac{\sinh(n-2r)\kappa(h+M)}{n-2r}\cos n\theta\right\}, \tag{14}$$

where h is the total mean depth

$$h = h_0+b, \tag{15}$$

and
$$M(\theta) = N(\theta)-b = \sum_1^{\infty} a_n\cos n\theta, \tag{16}$$
$$a_1 \equiv a.$$

The mean value of (14) over one period is

$$\mathscr{L} = \frac{1}{2\pi}\int_0^{2\pi}L\,d\theta = (\tfrac{1}{2}\beta^2-\gamma)h+\tfrac{1}{2}gb^2+\tfrac{1}{4}ga_1^2+\tfrac{1}{4}ga_2^2-\frac{\omega-\beta\kappa}{\kappa}(\mu_1A_1+\mu_2A_2)$$
$$+\kappa(\tfrac{1}{2}\mu_{11}A_1^2+\mu_{12}A_1A_2+\tfrac{1}{2}\mu_{22}A_2^2)+O(\epsilon^6), \tag{17}$$

where

$$\mu_1 = \tfrac{1}{2}\kappa a_1\cosh\kappa h+\tfrac{1}{4}\kappa^2a_1a_2\sinh\kappa h+\tfrac{1}{16}\kappa^3a_1^3\cosh\kappa h,$$
$$\mu_2 = \tfrac{1}{2}\kappa a_2\cosh 2\kappa h+\tfrac{1}{4}\kappa^2a_1^2\sinh 2\kappa h,$$
$$\mu_{11} = \tfrac{1}{4}\sinh 2\kappa h+\tfrac{1}{4}\kappa^2a_1^2\sinh 2\kappa h+\tfrac{1}{4}\kappa a_2,$$
$$\mu_{12} = \tfrac{1}{4}\kappa a_1\cosh 3\kappa h,$$
$$\mu_{22} = \tfrac{1}{8}\sinh 4\kappa h.$$

The variation with respect to A_1 and A_2 gives

$$\mathscr{L}_{A_1}\propto\mu_{11}A_1+\mu_{12}A_2-\frac{\omega-\beta\kappa}{\kappa^2}\mu_1 = 0,$$
$$\mathscr{L}_{A_2}\propto\mu_{12}A_1+\mu_{22}A_2-\frac{\omega-\beta\kappa}{\kappa^2}\mu_2 = 0.$$

With these relations for A_1 and A_2, the terms involving A_1, A_2 in (17) may be evaluated as

$$-\frac{1}{2}\frac{(\omega-\beta\kappa)^2}{\kappa^3}\frac{\mu_1^2\mu_{22}-2\mu_1\mu_2\mu_{12}+\mu_2^2\mu_{11}}{\mu_{11}\mu_{22}-\mu_{12}^2}. \tag{18}$$

Hence, the expression for $\mathscr{L}$ becomes

$$\mathscr{L} = (\tfrac{1}{2}\beta^2-\gamma)h+\tfrac{1}{2}gb^2+\tfrac{1}{4}ga_1^2+\tfrac{1}{4}ga_2^2$$
$$-\frac{1}{4}\frac{(\omega-\beta\kappa)^2}{\kappa T}\left\{a_1^2-\frac{2T^2-1}{4T^2}\kappa^2a_1^4-\frac{3-T^2}{2T}\kappa a_1^2a_2+(1+T^2)a_2^2\right\}+O(\epsilon^6), \tag{19}$$

where $T\equiv\tanh\kappa h$, and terms of order ϵ^6 have again been neglected.

The quantities (β, γ, b) and a_2 are $O(\epsilon^2)$ so that the lowest-order approximation, corresponding to the usual linear theory, is

$$\mathscr{L} = -\gamma h_0 + \tfrac{1}{4} g a_1^2 \left(1 - \frac{\omega^2}{g\kappa \tanh \kappa h_0}\right) + O(\epsilon^4). \tag{20}$$

(The term $-\gamma h_0$ does not contribute in the final equations and could be omitted.) Then, the variation with respect to a_1 gives the usual linear dispersion relation

$$\omega^2 = \omega_0^2(\kappa) + O(\epsilon^2), \tag{21}$$

where
$$\omega_0^2(\kappa) = g\kappa \tanh \kappa h_0. \tag{22}$$

In view of this, terms of order

$$\left(\frac{\omega^2}{\omega_0^2} - 1\right)\epsilon^4$$

will not be needed in (19) and it may be simplified to

$$\begin{aligned}\mathscr{L} = {} & (\tfrac{1}{2}\beta^2 - \gamma)\, h + \tfrac{1}{2} g b^2 + \tfrac{1}{4} g a_1^2 \left\{1 - \frac{(\omega - \beta\kappa)^2}{g\kappa \tanh \kappa h}\right\} \\ & + \tfrac{1}{4} g \left\{\frac{2T_0^2 - 1}{4T_0^2}\kappa^2 a_1^4 + \frac{3 - T_0^2}{2T_0}\kappa a_1^2 a_2 - T_0^2 a_2^2\right\} + O\left\{\epsilon^6, \left(\frac{\omega^2}{\omega_0^2} - 1\right)\epsilon^4\right\},\end{aligned} \tag{23}$$

where $T_0 = \tanh \kappa h_0$. The variation with respect to a_2 now gives the relation

$$a_2 = \{(3 - T_0^2)/4T_0^3\}\kappa a_1^2, \tag{24}$$

and (23) reduces to

$$\mathscr{L} = (\tfrac{1}{2}\beta^2 - \gamma)\, h + \tfrac{1}{2} g b^2 + \tfrac{1}{2} E \left\{1 - \frac{(\omega - \beta\kappa)^2}{g\kappa \tanh \kappa h}\right\} + \tfrac{1}{2} E^2 \frac{\kappa^2 D_0}{g \tanh \kappa h_0} + O\left\{\epsilon^6, \left(\frac{\omega^2}{\omega_0^2} - 1\right)\epsilon^4\right\}, \tag{25}$$

where
$$E = \tfrac{1}{2} g a_1^2 \equiv \tfrac{1}{2} g a^2, \tag{26}$$

and
$$D_0 = (9T_0^4 - 10T_0^2 + 9)/8T_0^3. \tag{27}$$

The expression in (25) is the final form of the Lagrangian as a function of the two triads (κ, ω, a) and (β, γ, b). It is often convenient to use E, which is proportional to the energy density, as a variable in place of a.

The variation with respect to a, or equivalently with respect to E, gives the non-linear dispersion relation

$$\frac{(\omega - \beta\kappa)^2}{g\kappa \tanh \kappa h} = 1 + \frac{2\kappa^2 D_0 E}{g \tanh \kappa h_0} + O(\epsilon^4). \tag{28}$$

The dependence on β and b (through $h = h_0 + b$) should be noticed, as well as the dependence on E.

3. Linear waves

For linear wave trains the average Lagrangian always takes the form

$$\mathscr{L} = G(\omega, \kappa)\, E,$$

where $G(\omega, \kappa) = 0$ is the dispersion relation (Whitham 1965*b*).

This is verified for the present case; in the lowest approximation,

$$\mathscr{L} = \frac{1}{2}\left(1 - \frac{\omega^2}{\omega_0^2(\kappa)}\right) E, \tag{29}$$

$$E = \tfrac{1}{2} g a^2, \quad \omega_0^2 = g\kappa \tanh \kappa h_0.$$

The dispersion of the wave train is then governed by the equations

$$-\frac{\partial \mathscr{L}_\omega}{\partial t} + \frac{\partial \mathscr{L}_\kappa}{\partial x} = 0, \quad \frac{\partial \kappa}{\partial t} + \frac{\partial \omega}{\partial x} = 0.$$

From (29), we have

$$\frac{\partial}{\partial t}\left(\frac{E}{\omega_0}\right) + \frac{\partial}{\partial x}\left(C_0 \frac{E}{\omega_0}\right) = 0, \tag{30}$$

$$\frac{\partial \kappa}{\partial t} + \frac{\partial \omega_0}{\partial x} = 0, \tag{31}$$

where C_0 is the linear group velocity

$$\left.\begin{aligned} C_0 &= \omega_0'(\kappa), \\ &= \tfrac{1}{2} c_0 \left(1 + \frac{\kappa h_0}{\sinh \kappa h_0 \cosh \kappa h_0}\right), \\ c_0 &= \omega_0/\kappa. \end{aligned}\right\} \tag{32}$$

(30) and (31) may be simplified to

$$\frac{\partial E}{\partial t} + C_0 \frac{\partial E}{\partial x} + E C_0' \frac{\partial \kappa}{\partial x} = 0, \tag{33}$$

$$\frac{\partial \kappa}{\partial t} + C_0 \frac{\partial \kappa}{\partial x} = 0. \tag{34}$$

A detailed discussion of these equations was given in the earlier references, but a brief résumé will help in the study of the non-linear case. The equations determine the laws of propagation of changes in wave-number κ and energy E with the group velocity C_0. Treated as a system, the pair (33) and (34) is not hyperbolic, since two independent combinations in characteristic form cannot be found. However, (34) can be solved first to give $\kappa =$ const. on the characteristics $dx/dt = C_0(\kappa)$, which are straight lines in the (x, t)-plane as a consequence. *With κ determined,* (33) can be written as a simple linear equation

$$dE/dt = -C_0'(\kappa)\, \kappa_x E \tag{35}$$

along the same characteristics. In this sense, the equations are hyperbolic and $C_0(\kappa)$ is a double characteristic velocity. (35) can be re-cast to show that the energy between two neighbouring characteristics remains constant with time.

As the next section will show, the effects of the small non-linearity of the basic wave train do not give merely a small correction to these results; the whole structure of the equations is different.

4. Non-linear waves

For non-linear waves, the second-order quantities (β, γ, b) are included and the full set of equations (7), (8), (9) is used with the Lagrangian given by (25). The dispersion relation follows from $\mathscr{L}_a = 0$ (or $\mathscr{L}_E = 0$) and it has been noted already in (28). It may be expanded, remembering that $h = h_0 + b$, to give

$$\omega = \omega_0 + \frac{\kappa^2 D_0}{c_0} E + \frac{\kappa B_0}{h_0} b + \kappa\beta + O(\epsilon^4), \tag{36}$$

where
$$B_0 = g h_0 \kappa / 2\omega_0 \cosh^2 \kappa h_0 = C_0 - \tfrac{1}{2} c_0. \tag{37}$$

In a similar way, the second functional relation $\mathscr{L}_b = 0$ gives

$$\gamma = g b + \frac{B_0}{c_0 h_0} E + O(\epsilon^4) \tag{38}$$

for the 'pseudo frequency' γ.

The other four derivatives of $\mathscr{L}$ are

$$\mathscr{L}_\gamma = -\frac{E}{\omega_0} + O(\epsilon^4), \quad \mathscr{L}_\kappa = C_0 \frac{E}{\omega_0} + O(\epsilon^4),$$
$$\mathscr{L}_\omega = -(h_0 + b) + O(\epsilon^4), \quad \mathscr{L}_\beta = \beta h_0 + (E/c_0) + O(\epsilon^4).$$

Substituting in (8), (9) and omitting error terms of order ϵ^4, we have

$$\frac{\partial}{\partial t}\left(\frac{E}{\omega_0}\right) + \frac{\partial}{\partial x}\left(\frac{C_0 E}{\omega_0}\right) = 0, \tag{39}$$

$$\frac{\partial b}{\partial t} + \frac{\partial}{\partial x}\left(\beta h_0 + \frac{E}{c_0}\right) = 0, \tag{40}$$

$$\frac{\partial \kappa}{\partial t} + \frac{\partial}{\partial x}\left(\omega_0 + \frac{\kappa^2 D_0}{c_0} E + \frac{\kappa B_0}{h_0} b + \kappa\beta\right) = 0, \tag{41}$$

$$\frac{\partial \beta}{\partial t} + \frac{\partial}{\partial x}\left(g b + \frac{B_0}{c_0 h_0} E\right) = 0. \tag{42}$$

To the same order of approximation, (39) can be replaced by

$$\frac{\partial E}{\partial t} + \frac{\partial}{\partial x}(C_0 E) = 0, \tag{43}$$

since
$$\frac{\partial \omega_0}{\partial t} + C_0 \frac{\partial \omega_0}{\partial x} = C_0\left(\frac{\partial \kappa}{\partial t} + C_0 \frac{\partial \kappa}{\partial x}\right) = O(\epsilon^2), \tag{44}$$

from (41).

These equations can be interpreted in physical terms. In (40), b and β measure the mean height and mean fluid velocity; hence, (40) is conservation of mass and it is seen that the additional transport due to the waves is E/c_0. (42) is related to the momentum equation but it is not quite in the correct form as it stands. The momentum density is the same as the mass flux, i.e.

$$\beta h_0 + (E/c_0).$$

Therefore, from (39), (42) and (44), the equation for conservation of momentum can be written

$$\frac{\partial}{\partial t}\left(\beta h_0 + \frac{E}{c_0}\right) + \frac{\partial}{\partial x}\left\{g h_0 b + (B_0 + C_0)\frac{E}{c_0}\right\} = 0.$$

The additional momentum flux due to the waves is

$$(B_0 + C_0)\frac{E}{c_0} = \left(2\frac{C_0}{c_0} - \tfrac{1}{2}\right)E.$$

(43) is clearly the energy equation for the waves. The total energy equation, including the mean flow etc., can also be derived. These expressions for mass flux, etc. have all been derived previously (Longuet-Higgins & Stewart 1960). In fact a more approximate version of (40)–(43) was discussed before (Whitham 1962); however, the crucial terms in ω beyond the linear approximation were not included, and the full significance of the equations was not really appreciated.

Deep-water waves

In the deep-water limit $\kappa h_0 \to \infty$, $c_0 \sim \sqrt{(g/\kappa)}$, $C_0 \sim \frac{1}{2}\sqrt{(g/\kappa)}$, $D_0 \to 1$, $B_0/c_0 \to 0$. It is then clear that the term $\kappa B_0 b/h_0$ in (41) may be neglected; its ratio to ω_0 is

$$\frac{B_0}{c_0}\frac{b}{h_0} \to 0, \quad \text{as} \quad \kappa h_0 \to \infty.$$

Since β is the mean velocity over the whole depth, one would certainly expect in this limit of infinite depth that $\beta \to 0$, and $\kappa\beta$ can also be neglected in (41). The orders of b and β can be estimated more carefully from (40) and (42). To a first approximation, changes in E propagate with the linear group velocity C_0. Therefore, in considering the resultant changes in b and β, the t-derivatives in (40) and (42) are C_0 times the corresponding x-derivatives. Thus,

$$-C_0 b + \beta h_0 + (E/c_0) = 0, \quad -C_0\beta + gb + (B_0/c_0 h_0)E = 0;$$

hence

$$b = -\frac{B_0 + C_0}{gh_0 - C_0^2}\frac{E}{c_0} \sim -\frac{1}{2}\frac{E}{gh_0} \quad \text{as} \quad \kappa h_0 \to \infty, \tag{45}$$

$$\beta = -\frac{E}{c_0 h_0} - \frac{C_0(B_0 + C_0)}{gh_0 - C_0^2}\frac{E}{c_0 h_0} \sim -\frac{E}{c_0 h_0} \quad \text{as} \quad \kappa h_0 \to \infty. \tag{46}$$

Perhaps the most important point is to note that the extra volume flux E/c_0, due to the waves, is balanced in (40) by a small mean counter-flow proportional to $1/h_0$, rather than by changes in the mean height b.

It is then clear that, in this limit, equations (41) and (43) for κ and E uncouple from the other pair and we have

$$\frac{\partial \kappa}{\partial t}+\frac{\partial}{\partial x}\left(\omega_0+\kappa^2\frac{E}{c_0}\right) = 0, \tag{47}$$

$$\frac{\partial E}{\partial t}+\frac{\partial}{\partial x}(C_0 E) = 0. \tag{48}$$

From (44), the E in (48) can always be replaced by E multiplied by any function of κ. Therefore, (47) and (48) are equivalent to

$$\kappa_t+C_0\kappa_x+(\kappa^2E/c_0)_x = 0, \tag{49}$$

$$\left(\frac{\kappa^2 E}{c_0}\right)_t+C_0\left(\frac{\kappa^2 E}{c_0}\right)_x+\left(\frac{\kappa^2 E}{c_0}\right)C_0'\kappa_x = 0. \tag{50}$$

The hyperbolic or elliptic nature of these equations may be established by searching for linear combinations that are in characteristic form. If m times (49) is added to (50), the resulting combination would be in characteristic form only if

$$m^2 = (\kappa^2E/c_0)\,C_0'.$$

Since $C_0' < 0$, m is imaginary and the equations are elliptic. The significance of this is discussed later.

General case of finite depth

In general, the four equations (40)–(43) are coupled together and the structure of the whole system must be considered. It is convenient to work with $\mathscr{E} = E/c_0$ in place of E; as noted before, by appeal to (44), (42) takes the same form in $\mathscr{E}$. Accordingly, the set of equations is written as

$$\kappa_t+C_0\kappa_x+\kappa^2D_0\mathscr{E}_x+(\kappa B_0/h_0)\,b_x+\kappa\beta_x+F\kappa_x = 0, \tag{51}$$

$$\mathscr{E}_t+C_0\mathscr{E}_x+\mathscr{E}C_0'\kappa_x = 0, \tag{52}$$

$$b_t+h_0\beta_x+\mathscr{E}_x = 0, \tag{53}$$

$$\beta_t+gb_x+\frac{B_0}{h_0}\mathscr{E}_x+\mathscr{E}\frac{B_0'}{h_0}\kappa_x = 0, \tag{54}$$

where F is a term of order $\mathscr{E}$. It turns out that it is both inconsistent and unnecessary to retain the term $F\kappa_x$ in (51). The operator on κ becomes

$$\partial/\partial t+(C_0+F)\,\partial/\partial x,$$

whereas a similar correction to the operator $\partial/\partial t+C_0\,\partial/\partial x$ on $\mathscr{E}$ in (52) would depend upon terms $\mathscr{E}\mathscr{E}_x$ which have been omitted. That this correction to the operator is unnecessary in both equations is seen as the analysis progresses. It is simplest to neglect the term at the outset and remember its effect as that of a term $O(\mathscr{E})$ added to C_0.

To find the characteristics of (51) to (54), linear combinations are considered in which multiples l_1, l_2, l_3 of (52), (53), (54), respectively, are added to (51). For the derivatives of κ, E, b, β to be in the same combination

$$\frac{\partial}{\partial t}+C\frac{\partial}{\partial x},$$

the multipliers l_i and the velocity C must satisfy

$$\begin{aligned} C &= C_0+l_1 C_0'\mathscr{E}+l_3(B_0'/h_0)\mathscr{E}, \\ l_1 C &= \kappa^2 D_0+l_1 C_0+l_2+l_3(B_0/h_0), \\ l_2 C &= (\kappa B_0/h_0)+l_3 g, \\ l_3 C &= \kappa+l_2 h_0. \end{aligned}$$

The errors are factors $1+O(\mathscr{E})$ in every term displayed. Solving this set of algebraic equations, we have

$$\left.\begin{aligned} l_1 &= \frac{C-C_0}{C_0'\mathscr{E}}+\frac{\kappa}{h_0}\frac{B_0'}{C_0'}\frac{B_0+C}{gh_0-C^2}, \\ l_2 &= -\frac{\kappa}{h_0}\frac{gh_0+B_0 C}{gh_0-C^2}, \quad l_3 = -\kappa\frac{B_0+C}{gh_0-C^2}, \end{aligned}\right\} \tag{55}$$

and

$$(C-C_0)^2+\frac{\kappa}{h_0}B_0'\frac{B_0+C}{gh_0-C^2}(C-C_0)\mathscr{E} = \left\{\kappa h_0 D_0-\frac{gh_0+2B_0 C_0+B_0^2}{gh_0-C^2}\right\}\frac{\kappa C_0'}{h_0}\mathscr{E}. \tag{56}$$

The errors in the equation for C are proportional to $\mathscr{E}(C-C_0)$ and $\mathscr{E}^2$. These errors take care of the F term, as may be seen by noting the effect of a term proportional to $\mathscr{E}$ added to C_0. The second term in (56) is retained even though it is proportional to $(C-C_0)\mathscr{E}$, since roots close to $\pm\sqrt{(gh_0)}$ occur.

Two of the roots of (56) tend to C_0 as $\mathscr{E}\to 0$ and correspond to the double root C_0 of the linear wave theory. For these two roots, the term proportional to $(C-C_0)\mathscr{E}$ can be neglected with the errors, and the roots are given approximately by

$$C = C_0\pm\left[\frac{\kappa C_0'}{h_0}\mathscr{E}\left\{\kappa h_0 D_0-\frac{gh_0+2B_0 C_0+B_0^2}{gh_0-C_0^2}\right\}\right]^{\frac{1}{2}}. \tag{57}$$

Since $C_0' < 0$, these roots are real if

$$(gh_0+2B_0 C_0+B_0^2)/(gh_0-C_0^2) > \kappa h_0 D_0. \tag{58}$$

The terms on the left of this inequality come from the coupling with mean height and velocity, b and β, while the term on the right comes from the non-linear dependence of the frequency on the amplitude. If the roots given in (57) are real, the double characteristic of linear waves splits into two distinct characteristics with velocities which differ from C_0 by an amount proportional to the amplitude $\sqrt{(2E/g)}$. The fact that the correction is proportional to $\sqrt{E}$ rather than E itself is due to the double root in the limit $E\to 0$. This also explains why the term F in (51) could be neglected.

The other two roots of (56) tend to $\pm\sqrt{(gh_0)}$ as $\mathscr{E}\to 0$, and correspond to the characteristics of simple shallow-water theory. The correction proportional to $\mathscr{E}$ can be determined from (56), but it suffices for the present discussion to know that these roots are real.

The inequality (58) determines the type of the system. If (58) holds, the system is hyperbolic; otherwise it has an elliptic part. The quantities in (58) are defined in (27), (32) and (37). The dependence on gh_0 cancels out, and the inequality depends on the value of κh_0 only. As $\kappa h_0 \to \infty$, the left-hand side tends to 1 while the right-hand side is asymptotic to κh_0; hence the equations are elliptic for sufficiently large values of κh_0. As $\kappa h_0 \to 0$, $\kappa h_0 D_0 \sim 9/(8\kappa^2 h_0^2)$ while the left-hand side of (58) is asymptotic to $9/(4\kappa^2 h_0^2)$; hence, the equations are hyperbolic for sufficiently small values of κh_0. The critical value is found to be $\kappa h_0 = 1{\cdot}36$.

5. Significance of the type of the averaged equations

If the averaged equations are hyperbolic, changes in the wave-train parameters propagate according to the usual theory of characteristics. The characteristic velocities provide a generalization of the linear group velocity to non-linear problems. In the non-linear theory, the splitting of the double characteristics $C = C_0$ into two separate families of characteristics introduces important changes in the nature of the solutions. This has been discussed in detail in previous papers.

When the equations have an elliptic part, there is a certain kind of instability in the wave train. Consider a general system of quasi-linear equations:

$$U_t + AU_x = 0,$$

where U is a column vector, A is the coefficient matrix with elements depending upon U. Our averaged equations are of this form. Now, consider a small perturbation of U on a constant solution U_0. The linearized equation for the perturbation $u = U - U_0$ is

$$u_t + A_0 u_x = 0, \tag{59}$$

where $A_0 = A(U_0)$. Elementary solutions of (59) can be found in the form

$$u = u_0 e^{i\mu(x-Ct)}. \tag{60}$$

This is a solution of (59) provided

$$(A_0 - CI)\, u_0 = 0,$$

where I is the unit matrix. The condition for non-trivial solutions is

$$\det |A_0 - CI| = 0. \tag{61}$$

The eigenvalues C are just the characteristic velocities of the system. If they are real, the solution (60) remains a small perturbation. If there are imaginary roots, corresponding to an elliptic part, some terms in the perturbation solution (60) will have exponential growth with time. Thus a periodic wave train will ultimately break up into some other form.

For water waves, therefore, a periodic wave train should be unstable in this sense when $\kappa h_0 > 1{\cdot}36$.

6. Relation with Benjamin's approach

Benjamin (1966) found the instability result for the deep-water limit. In his approach, the analysis depends upon a subtle kind of resonance between two small modulations with wave-number $\kappa \pm \mu$ and the main wave κ. He starts from a solution of the form

$$\phi = \alpha(t)\, e^{i\kappa x} + \alpha^*(t)\, e^{-i\kappa x} + \alpha_+(t)\, e^{i(\kappa+\mu)x} + \alpha_+^*(t)\, e^{-i(\kappa+\mu)x} + \alpha_-(t)\, e^{i(\kappa-\mu)x} + \alpha_-^*(t)\, e^{-i(\kappa-\mu)x}, \tag{62}$$

where $\alpha_+(t)$ and $\alpha_-(t)$ are small compared with $\alpha(t)$. (The asterisk denotes complex conjugates.) In the linear theory, this would be a solution if

$$\alpha(t) = \alpha(0)\, e^{-i\omega_0(\kappa)t}, \quad \alpha_\pm(t) = \alpha_\pm(0)\, e^{-i\omega_0(\kappa\pm\mu)t}.$$

However, if the linear solution is taken as the first term in a naïve expansion in amplitude, a resonance in the higher-order cubic terms produces terms with denominators proportional to μ, and the expansion is not uniformly valid as $\mu \to 0$. The resonance arises in the cubic terms because, for example, the product of $(\alpha\, e^{i\kappa x})^2$ and $\alpha_+^*\, e^{-i(\kappa+\mu)x}$ produces a forcing term $\alpha^2\alpha_+^*\, e^{i(\kappa-\mu)x}$ which has the same wave-number as one of the terms in (62). To remedy this, the coefficients $\alpha(t), \alpha_\pm(t)$ must be determined in a way that includes such higher-order resonant terms, which repeat the basic wave-numbers $\pm\kappa$, $\pm(\kappa \pm \mu)$. The revised coefficients $\alpha_\pm(t)$ may still be oscillatory, indicating stability, or have an exponential growth, indicating instability and transfer of energy to the side bands $\kappa \pm \mu$.

The relation with the present analysis is that (62) can be written as a slowly varying wave train in the case $\mu \ll \kappa$. For consider a slowly varying wave train

$$\phi = \tfrac{1}{2}a\, e^{i\theta} + \tfrac{1}{2}a^*\, e^{-i\theta}, \tag{63}$$

where a, θ_x, θ_t are slowly varying functions. Suppose in particular that

$$a = a_0 + a_1, \quad \theta = \theta_0 + \theta_1,$$

where a_0 is constant, $\theta_0 = \kappa x - \omega t$ and $a_1/a_0 \ll 1$, $\theta_1/\theta_0 \ll 1$. That is, $a_0\, e^{i\theta_0}$ is a basic wave train and a_1, θ_1 represent a small modulation both in amplitude and phase. The basic wave train has small enough amplitude to keep the sinusoidal form but the dispersion relation is non-linear: ω depends on a_0 as well as κ. Now expand (63) as

$$\phi = a_0\, e^{i\theta_0} + a_1\, e^{i\theta_0} + i\theta_1\, a_0\, e^{i\theta_0} + \text{conjugate}, \tag{64}$$

and take

$$a_1 = A_+\, e^{i\mu x} + A_-\, e^{-i\mu x}, \quad \theta_1 = \Theta_+\, e^{i\mu x} + \Theta_-\, e^{-i\mu x}. \tag{65}$$

Since $\theta_0 = \kappa x - \omega t$, it is clear that (64) can then be put in the form (62) and vice versa. The subtle interplay between the two side bands $\kappa+\mu$ and $\kappa-\mu$ in Benjamin's approach appears more straightforward here and is included automatically by the coupling of the amplitude and phase changes. The functions in (65) are slowly varying provided $\mu \ll \kappa$.

When the details are followed through, the results of the two theories agree for $\mu \ll \kappa$. As noted already, Benjamin's theory does not require $\mu \ll \kappa$. On the other hand, the present theory would go through if the basic wave train were highly non-linear and could not be approximated by the sinusoidal form.

The comparison of the methods has been given for the deep-water case in which changes in the mean height and mean velocity can be ignored. The comparison for the more general case is similar.

This research was supported by the Office of Naval Research under contract Nonr-220(56). Reproduction in whole or in part is permitted for any purpose of the United States Government.

REFERENCES

Bateman, H. 1944 *Partial Differential Equations*. New York: Dover.
Benjamin, T. B. 1966 *J. Fluid Mech.* **25**, 241.
Lighthill, M. J. 1965 *J.I.M.A.* **1**, 269.
Longuet-Higgins, M. S. & Stewart, R. W. 1960 *J. Fluid Mech.* **8**, 565.
Luke, J. C. 1966 *J. Fluid Mech.* **27**, 395.
Whitham, G. B. 1962 *J. Fluid Mech.* **12**, 135.
Whitham, G. B. 1965*a* *Proc. Roy. Soc.* A, **283**, 238.
Whitham, G. B. 1965*b* *J. Fluid Mech.* **22**, 273.