J. Fluid Mech. (1970), *vol.* 44, *part* 2, *pp.* 373–395
Printed in Great Britain

Two-timing, variational principles and waves

By G. B. WHITHAM
California Institute of Technology

(Received 10 January 1970)

In this paper, it is shown how the author's general theory of slowly varying wave trains may be derived as the first term in a formal perturbation expansion. In its most effective form, the perturbation procedure is applied directly to the governing variational principle and an averaged variational principle is established directly. This novel use of a perturbation method may have value outside the class of wave problems considered here. Various useful manipulations of the average Lagrangian are shown to be similar to the transformations leading to Hamilton's equations in mechanics. The methods developed here for waves may also be used on the older problems of adiabatic invariants in mechanics, and they provide a different treatment; the typical problem of central orbits is included in the examples.

1. Introduction

The purpose of this paper is to show how the variational theory for slowly varying non-linear wave trains (Whitham 1965*a*, *b*, 1967*a*, *b*) may be formally justified as the first term in a consistent perturbation expansion. The theory referred to uses an 'averaged' variational principle, and the derivation of the equations governing the slow variations of a nearly periodic wave train is then surprisingly general and simple. The averaging procedure is intuitively correct, but it is not immediately clear how the averaged variational principle would appear as the first term of a detailed expansion. Luke (1966) has shown how the final results can be justified in such a scheme, but he works with the Euler equations rather than directly on the variational principle. This has the disadvantages that the simplicity and conciseness of the variational technique are lost and the formal justification no longer follows the intuitive derivation. It does, however, justify the results and has its own points of interest; it can be extended to dissipative systems, for example, where a variational principle may not be known.

It has now been seen how to apply the 'two-timing' approach used by Luke directly on the variational principle. This is the main topic of the paper. A second topic concerns some subtleties in the actual use of the variational principle. Various equivalent forms of the averaged variational principle can be introduced, and one of them has enormous advantages over the others. In specific problems, the optimum one can usually be found without too much difficulty. However, there are subtleties involved which were not clearly recognized in the

original versions. These are now explained, and general procedures are given for finding the preferred form.

The original theory is reviewed briefly in the next section, and the perturbation schemes are then taken up in §§ 3 to 6. The questions regarding manipulations of the form of the averaged variational principle are discussed in § 7, with illustrative examples in § 8.

The justification of the averaged variational principle in §§ 5 and 6 is the result of joint discussions with Mr M. E. Delaney.

It is a great pleasure to contribute to this volume, and in this way acknowledge a long-standing and continuing debt to Sydney Goldstein. Both as teacher and friend, his help and encouragement have always been invaluable.

2. The averaged variational principle

In all problems where the equations admit uniform periodic wave trains as solutions it appears to be generally true that the system can be derived from a variational principle. The simplest case occurs when there is one dependent variable $u(\mathbf{x}, t)$ and the variational principle takes the form

$$\delta \iint L(u_t, u_{\mathbf{x}}, u)\, dt\, d\mathbf{x} = 0. \tag{1}$$

The Euler equation is

$$\frac{\partial L_1}{\partial t} + \frac{\partial L_2}{\partial \mathbf{x}} - L_3 = 0, \tag{2}$$

where

$$L_1 = \frac{\partial L}{\partial u_t}, \quad L_2 = \frac{\partial L}{\partial u_{\mathbf{x}}}, \quad L_3 = \frac{\partial L}{\partial u}. \tag{3}$$

If $\mathbf{x}$ is a vector with components x_i, the quantities $u_{\mathbf{x}}$ and L_2 are vectors with components

$$\partial u / \partial x_i, \quad \partial L / \partial u_{x_i},$$

respectively, and the second term in (2) is the divergence

$$\frac{\partial L_2}{\partial \mathbf{x}} = \frac{\partial}{\partial x_i}\left(\frac{\partial L}{\partial u_{x_i}}\right).$$

We are concerned with dispersive wave problems in which the equation (2) for u has special solutions

$$u = U_0(\theta), \quad \theta = \kappa_i x_i - \omega t,$$

where $\boldsymbol{\kappa}$, ω are constants and where $U_0(\theta)$ turns out to be a periodic function of θ. Then $\boldsymbol{\kappa}$, ω are the wave-number and frequency, respectively. Since (2) is a second-order equation in u, there will be two constants of integration. One will determine the amplitude, a; the other will be an arbitrary shift in the phase θ and may be omitted here. The three parameters $\omega, \boldsymbol{\kappa}, a$ will not be independent, but must satisfy a 'dispersion relation'

$$G(\omega, \boldsymbol{\kappa}, a) = 0, \tag{4}$$

if (2) is to be satisfied. In linear problems

$$U_0 = a \cos \theta, \tag{5}$$

and the dispersion relation (4) does not involve the amplitude a.

For slowly varying wave trains, it is argued that the form of the solution, $u = U_0(\theta, a)$ is maintained, but a will not be constant nor will θ be linear in $\mathbf{x}$ and t. The wave-number $\boldsymbol{\kappa}$ and frequency ω are generalized by defining them as

$$\omega = -\theta_t, \quad \kappa_i = \theta_{x_i}. \tag{6}$$

The parameters ω, $\boldsymbol{\kappa}$, a will now be slowly varying functions of $\mathbf{x}$ and t corresponding to the slow modulation of the wave train. The aim is to derive equations for them.

If the period of the function $u = U_0(\theta, a)$ is normalized to 2π, the averaged Lagrangian is defined to be

$$\mathscr{L}(\omega, \boldsymbol{\kappa}, a) = \frac{1}{2\pi}\int_0^{2\pi} L\, d\theta, \tag{7}$$

and is calculated by substituting the *uniform* periodic solution $u = U_0(\theta, a)$ in L. The dependence on $\omega, \boldsymbol{\kappa}$ comes from introducing $u_t = -\omega U_0'$, $u_{\mathbf{x}} = \boldsymbol{\kappa} U_0'$ into $\mathscr{L}$; the integration with respect to θ is carried out holding ω, $\boldsymbol{\kappa}$, a constant. The claim is, then, that the 'averaged' variational principle holds and that the equations for $(\omega, \boldsymbol{\kappa}, a)$ follow from

$$\delta \iint \mathscr{L}(\omega, \boldsymbol{\kappa}, a)\, dt\, d\mathbf{x} = 0. \tag{8}$$

The quantities ω and $\boldsymbol{\kappa}$ are defined by (6) in terms of the phase function θ, so we must derive the consequences of (8) for independent variations $\delta\theta$ and δa. From the variation δa we have

$$\mathscr{L}_a(\omega, \boldsymbol{\kappa}, a) = 0, \tag{9}$$

and from the variation $\delta\theta$ we have

$$\frac{\partial}{\partial t}\mathscr{L}_\omega - \frac{\partial}{\partial x_i}\mathscr{L}_{\kappa_i} = 0. \tag{10}$$

Equation (9) is a functional relation between $(\omega, \boldsymbol{\kappa}, a)$, and it can only be the dispersion relation (4). One can use (10) as a second-order equation in θ, or work with $\omega, \boldsymbol{\kappa}$, and supplement (10) by the consistency equations

$$\frac{\partial \kappa_i}{\partial t} + \frac{\partial \omega}{\partial x_i} = 0, \quad \frac{\partial \kappa_i}{\partial x_j} - \frac{\partial \kappa_j}{\partial x_i} = 0, \tag{11}$$

derived from (6). The latter course is usually preferable.

The study of the consequences of (9)–(11), the interpretation of $\mathscr{L}_\omega$, $\mathscr{L}_{\kappa_i}$ as 'adiabatic invariants', the use of these equations to extend the concept of group velocity to non-linear problems, the relation of the stability of the periodic wave with the type of the equations, as well as extensions of the theory to cases with more than one dependent variable and to non-homogeneous media, are all given in the previous papers.

Two special cases will be useful in the present discussion. First, in any linear problem, L will be quadratic in u. Hence, when

$$U_0 = a\cos\theta$$

is substituted in (7), $\mathscr{L}$ *must* take the form

$$\mathscr{L}(\omega, \boldsymbol{\kappa}, a) = F(\omega, \boldsymbol{\kappa})a^2. \tag{12}$$

Quite generally, for linear problems with higher-order derivatives of u or with

more dependent variables u, the same argument goes through. The Lagrangian L must be quadratic in the variables u, and, as a consequence, $\mathscr{L}$ must be quadratic in a and take the form given in (12). Then the dispersion relation (9) becomes

$$F(\omega, \boldsymbol{\kappa}) = 0. \tag{13}$$

Hence, without detailed calculation, the function $F(\omega, \boldsymbol{\kappa})$ in (12) *must* be the dispersion function for the problem. We also see in general that the stationary value for $\mathscr{L}$ is zero in linear problems. In simple cases where L is the difference between kinetic and potential energy, this proves the well-known result that the average kinetic and potential energies are equal.

It should be noted that the dispersion relation between ω and $\boldsymbol{\kappa}$ is not used in the course of the calculation of $\mathscr{L}$. The result in (12) shows clearly that this is crucial. If the dispersion relation were used, the class of functions would be too restrictive and we would obtain merely the final stationary value $\mathscr{L} = 0$. In non-linear problems, this separation of the dispersion relation from the *form* of the periodic solution is more complicated. It is discussed in general in §7.

The second special case is a simple example to illustrate these non-linear effects. It is also a useful model to focus the general discussion of the perturbation methods. It is a non-linear version of the Klein–Gordon equation, namely

$$u_{tt} - u_{xx} + V'(u) = 0.$$

The Lagrangian is
$$L = \tfrac{1}{2}u_t^2 - \tfrac{1}{2}u_x^2 - V(u). \tag{14}$$

The periodic solution $u = U_0(\theta)$ satisfies

$$(\omega^2 - \kappa^2)U_0''(\theta) + V'(U_0) = 0, \tag{15}$$

which has the integral
$$\tfrac{1}{2}(\omega^2 - \kappa^2)U_0'^2 + V(U_0) = A. \tag{16}$$

The constant of integration A can be used as a parameter equivalent to the amplitude a. Equation (16) can be solved to give the function $U_0(\theta)$ in the inverse form

$$\theta = [\tfrac{1}{2}(\omega^2 - \kappa^2)]^{\frac{1}{2}} \int \frac{dU_0}{[A - V(U_0)]^{\frac{1}{2}}}. \tag{17}$$

Since θ is normalized so that the period is 2π, it follows that

$$2\pi = [\tfrac{1}{2}(\omega^2 - \kappa^2)]^{\frac{1}{2}} \oint \frac{dU_0}{[A - V(U_0)]^{\frac{1}{2}}}, \tag{18}$$

where $\oint$ denotes integration over a complete period. This is the dispersion relation between ω, κ and A.

The averaged Lagrangian is first introduced directly from (14) as

$$\bar{L} = \frac{1}{2\pi}\int_0^{2\pi} \{\tfrac{1}{2}(\omega^2 - \kappa^2)U_0'^2 - V(U_0)\}\, d\theta. \tag{19}$$

We then use (16) to obtain a function of ω, κ, A in the following way. We have successively

$$\begin{aligned} \bar{L} &= \frac{1}{2\pi}\int_0^{2\pi} (\omega^2 - \kappa^2)U_0'^2\, d\theta - A \\ &= \frac{1}{2\pi}\int_0^{2\pi} (\omega^2 - \kappa^2)U_0'\, dU_0 - A \\ &= \frac{1}{2\pi}[2(\omega^2 - \kappa^2)]^{\frac{1}{2}} \oint [A - V(U_0)]^{\frac{1}{2}}\, dU_0 - A. \end{aligned} \tag{20}$$

In this final form the U_0 has become merely a dummy variable of integration and we have a definite function of ω, κ, A. The symbol $\mathscr{L}(\omega, \kappa, A)$ is now reserved for this final form. The function $\mathscr{L}(\omega, \kappa, A)$ is used in (9) and (10). It is observed immediately that the dispersion relation (18) is just $\mathscr{L}_A = 0$, in accordance with the general result.

The important point about the manipulation in (20) is that only the equation (16) is used. Although, as noted in (17) and (18), this equation contains the solution for U_0 and the dispersion relation, these are not used explicitly. If they were, it would be possible to restrict the averaged Lagrangian too much to use its variation. Then (9) and (10) would not follow. For example, the dispersion relation (18) must not be used to eliminate $\omega^2 - \kappa^2$ from (20) altogether. We shall see later, in §7, how to describe the manipulation in general terms and how to justify the results.

3. Two-timing on the Euler equations

The so-called 'two-timing' method recognizes explicitly in the expressions for the dependent variables that changes are occurring on two time scales: the 'fast' oscillations of the wave train and the 'slow' variations of the parameters $(\omega, \boldsymbol{\kappa}, a)$. There are two corresponding length scales. The method originated in the study of the ordinary differential equations governing the non-linear vibrations of mechanical systems, where it was apparently first introduced by Krylov and Bogoliubov. It has received considerable development recently (see Cole 1968). Luke (1966) adapted it to the partial differential equations of wave problems using a basic paper by Kuzmak (1959) as a source.

The main idea is to express $u(\mathbf{x}, t)$ in the form

$$u(\mathbf{x}, t) = U(\theta, \mathbf{X}, T, \epsilon), \tag{21}$$

where

$$\theta = \epsilon^{-1}\Theta(\mathbf{X}, T), \quad \mathbf{X} = \epsilon\mathbf{x}, \quad T = \epsilon t, \tag{22}$$

and the small parameter ϵ measures the ratio of the fast time scale to the slow time scale. If the wave-number $\boldsymbol{\kappa}$ and frequency ω are introduced, we have

$$\left.\begin{aligned} \omega(\mathbf{X}, T) &= -\theta_t = -\Theta_T, \\ \boldsymbol{\kappa}(\mathbf{X}, T) &= \theta_{\mathbf{x}} = \Theta_{\mathbf{X}}. \end{aligned}\right\} \tag{23}$$

The scaling has been arranged so that

$$\frac{\partial\omega}{\partial t} = \epsilon\frac{\partial\omega}{\partial T}, \quad \frac{\partial\omega}{\partial x_i} = \epsilon\frac{\partial\omega}{\partial X_i}, \tag{24}$$

with similar expressions for $\boldsymbol{\kappa}$, and

$$\frac{\partial u}{\partial t} = -\omega\frac{\partial U}{\partial\theta} + \epsilon\frac{\partial U}{\partial T}, \quad \frac{\partial u}{\partial x_i} = \kappa_i\frac{\partial U}{\partial\theta} + \epsilon\frac{\partial U}{\partial X_i}. \tag{25}$$

If $\mathbf{X}$, T, ω, $\boldsymbol{\kappa}$, U are all taken to be $O(1)$ quantities, the scaling has been arranged so that ω, $\boldsymbol{\kappa}$ are slowly varying quantities, and so that u has a slow variation in addition to its oscillation with the phase θ.

The basis of the technique is to use the function $U(\theta, \mathbf{X}, T, \epsilon)$ explicitly as a function of all the independent variables θ, $\mathbf{X}$, T even though ultimately the extra variable θ is to be related back to $\mathbf{X}$ and T by (22). This extra freedom can be used to advantage in obtaining a uniformly valid expansion for u.

The relevant differential equation in the present case is the Euler equation (2). For simplicity in the discussion, the case of one space dimension x is considered; the extension is trivial. The first step is to convert the equation (2) into an equation for† $U(\theta, X, T)$. In doing this, we must use $\theta = \epsilon^{-1}\Theta(X, T)$. As in (25), t and x derivatives become

$$\frac{\partial}{\partial t} = \Theta_T \frac{\partial}{\partial \theta} + \epsilon \frac{\partial}{\partial T}, \quad \frac{\partial}{\partial x} = \Theta_X \frac{\partial}{\partial \theta} + \epsilon \frac{\partial}{\partial X}. \tag{26}$$

To preserve the symmetry between x and t it will be convenient to work with

$$\nu = \Theta_T = -\omega$$

rather than ω. Then, (2) becomes

$$\frac{\partial}{\partial \theta}(\nu L_1 + \kappa L_2) - L_3 + \epsilon \frac{\partial L_1}{\partial T} + \epsilon \frac{\partial L_2}{\partial X} = 0. \tag{27}$$

In L_1, L_2, L_3, which are defined in terms of L by (3), the arguments are as shown:

$$L_j = L_j(\nu U_\theta + \epsilon U_T,\ \kappa U_\theta + \epsilon U_X,\ U). \tag{28}$$

At this point, the relation of θ to X, T may be dropped temporarily, and (27) can be considered as an equation for the function $U(\theta, X, T)$ of the three *independent* variables θ, X, T. Clearly, if this function is found, then $U(\epsilon^{-1}\Theta, X, T)$ solves the original problem. It should be noted that $\Theta(X, T)$ still appears in (27), through $\nu = \Theta_T$, $\kappa = \Theta_X$, but the relation of Θ to the argument θ in U is dropped.

The solution of (27) is now obtained by expanding U formally in a power series

$$U(\theta, X, T, \epsilon) = \sum_{n=0}^{\infty} \epsilon^n U_n(\theta, X, T),$$

and equating the terms of successive orders in ϵ to zero. To lowest order in ϵ, only U_0 is involved and we have

$$\frac{\partial}{\partial \theta}\{\nu L_1^{(0)} + \kappa L_2^{(0)}\} - L_3^{(0)} = 0, \tag{29}$$

where

$$L_j^{(0)} = L_j(\nu U_{0\theta},\ \kappa U_{0\theta},\ U_0). \tag{30}$$

Since (29), (30) involve only θ derivatives of U_0, this is effectively an ordinary differential equation, in fact that of the uniform periodic wave train. In solving it an arbitrary 'constant' A will arise which is equivalent to the amplitude a. Of course, this 'constant' must now be allowed to be a function of (X, T). Just the amplitude modulation required!

At this lowest order, U_0 will be determined in its dependence on θ, but will involve the three parameters ν, κ, A which are all functions of X, T. There is one

† The dependence of U on the small parameter ϵ will not be shown, whenever it is not a point to be emphasized.

relation between them at this stage. In solving for U_0 and normalizing the period in θ to 2π, the parameters ν, κ, A will be required to satisfy the dispersion relation; in this direct attack on the Euler equations there are no subtleties about whether the dispersion relation is or is not used. In the Klein–Gordon example (14), the lowest-order equation (29) is just (15). It integrates to (16), now with $A = A(X, T)$, the dependence of U_0 on θ is determined, together with the dispersion relation as in (17), (18).

Further relations between ν, κ, A are obtained by going to the next order in ϵ. The next-order terms in (27) involve U_1, U_0, ν, κ. Only θ derivatives of U_1 occur, so it is again, effectively, an ordinary differential equation for U_1. The solution will not be periodic in θ, there will be 'secular' terms linear in θ, unless an appropriate condition is enforced on the non-homogeneous terms. For the expansion of U to be uniformly valid in θ, the secular terms *must* be suppressed. We shall call this requirement a 'secular condition'. It is similar to the orthogonality condition required of the non-homogeneous terms in certain linear problems of eigenfunction expansions. The secular condition gives a further equation for ν, κ, A.

The suppression of secular terms is most easily carried out on conservation equations. Indeed the whole two-timing procedure is considerably simplified by working with conservation equations, and some of the complication Luke ran into is avoided.

The first step is to write (27) in the equivalent conservation form†

$$\frac{\partial}{\partial\theta}\{(\nu L_1 + \kappa L_2)U_\theta - L\} + \epsilon\frac{\partial}{\partial T}(U_\theta L_1) + \epsilon\frac{\partial}{\partial X}(U_\theta L_2) = 0. \tag{31}$$

We shall denote this by
$$\frac{\partial R}{\partial\theta} + \epsilon\frac{\partial P}{\partial T} + \epsilon\frac{\partial Q}{\partial X} = 0 \tag{32}$$

for ease of discussion. The quantities P, Q, R are functions of U, U_θ, U_X, U_T, ν, κ and ϵ. If we now expand U in a power series in ϵ, the quantities P, Q, R will have corresponding expansions

$$P = \sum_0^\infty P^{(n)}\epsilon^n, \text{ etc.}$$

The lowest terms $P^{(0)}$, $Q^{(0)}$, $R^{(0)}$ depend on U_0, ν, κ; the next terms $P^{(1)}$, $Q^{(1)}$, $R^{(1)}$ include also U_1.

The first two terms of (32) are
$$\frac{\partial R^{(0)}}{\partial\theta} = 0, \tag{33}$$

$$\frac{\partial R^{(1)}}{\partial\theta} = -\frac{\partial P^{(0)}}{\partial T} - \frac{\partial Q^{(0)}}{\partial X}. \tag{34}$$

The first one is equivalent to (29) and leads to the immediate first integral

$$R^{(0)} \equiv (\nu L_1^{(0)} + \kappa L_2^{(0)})U_{0\theta} - L^{(0)} = A(X, T). \tag{35}$$

(This is (16) for the Klein–Gordon example.) The second equation (34) has to be solved for U_1, which occurs in $R^{(1)}$. A solution uniformly valid in θ requires U, and

† A natural derivation of this equation, free of ingenuity, will be noted later in §5.

hence each U_n, to be *periodic* in θ with period 2π. This is only possible in (34) if the integral of the right-hand side with respect to θ over one period is zero. That is, to avoid secular terms, we must demand

$$\frac{\partial}{\partial T}\left(\frac{1}{2\pi}\int_0^{2\pi} P^{(0)}d\theta\right)+\frac{\partial}{\partial X}\left(\frac{1}{2\pi}\int_0^{2\pi} Q^{(0)}d\theta\right) = 0. \tag{36}$$

From the definitions of P, Q, R (see (31) and (32)), it is

$$\frac{\partial}{\partial T}\left(\frac{1}{2\pi}\int_0^{2\pi} U_{0\theta}L_1^{(0)}d\theta\right)+\frac{\partial}{\partial X}\left(\frac{1}{2\pi}\int_0^{2\pi} U_{0\theta}L_2^{(0)}d\theta\right) = 0. \tag{37}$$

This is the remaining equation for ν, κ, A.

In this approach, (35) is solved completely for U_0 and the dispersion relation. The results are substituted in (37). The tie-in with the averaged Lagrangian raises the more subtle question of just how (35) is to be used in conjunction with (37). We may note, however, that the averaged Lagrangian is

$$\overline{L^{(0)}} = \frac{1}{2\pi}\int_0^{2\pi} L(\nu U_{0\theta}, \kappa U_{0\theta}, U_0)\,d\theta. \tag{38}$$

If derivatives are taken, keeping U_0 fixed, then (37) can indeed be written

$$\frac{\partial}{\partial T}\overline{L_\nu^{(0)}}+\frac{\partial}{\partial X}\overline{L_\kappa^{(0)}} = 0. \tag{39}$$

The form $\mathscr{L}(\nu, \kappa, A)$, with appropriate careful use of (35) in (38), still gives

$$\frac{\partial}{\partial T}\mathscr{L}_\nu+\frac{\partial}{\partial X}\mathscr{L}_\kappa = 0, \tag{40}$$

even though further dependence on ν and κ has been introduced into (38) via U_0. Moreover, the dispersion relation becomes $\mathscr{L}_A = 0$. This preferred form is derived in §7 after the averaged variational principle has been justified and used directly to derive the results of this section.

For the Klein–Gordon example, L is given by (14), so that

$$L^{(0)} = \tfrac{1}{2}(\nu^2-\kappa^2)U_{0\theta}^2 - V(U_0)$$

and (39) becomes

$$\frac{\partial}{\partial T}\left(\frac{\nu}{2\pi}\int_0^{2\pi} U_{0\theta}^2 d\theta\right)-\frac{\partial}{\partial X}\left(\frac{\kappa}{2\pi}\int_0^{2\pi} U_{0\theta}^2 d\theta\right) = 0. \tag{41}$$

On the other hand, from (20), $\mathscr{L}$ is

$$\mathscr{L} = \frac{1}{2\pi}[2(\nu^2-\kappa^2)]^{\frac{1}{2}}\oint[A-V(U_0)]^{\frac{1}{2}}dU_0 - A, \tag{42}$$

and (40) becomes

$$\frac{\partial}{\partial T}\left(\frac{1}{2\pi}\frac{\nu}{[\nu^2-\kappa^2]^{\frac{1}{2}}}\oint[A-V(U_0)]^{\frac{1}{2}}dU_0\right)-\frac{\partial}{\partial X}\left\{\frac{1}{2\pi}\frac{\kappa}{[\nu^2-\kappa^2]^{\frac{1}{2}}}\oint[A-V(U_0]^{\frac{1}{2}}dU_0\right\} = 0. \tag{43}$$

The equivalence of these two forms of the secular condition, (41) and (43), is easily established for this example from (16).

Two-timing on the equations, then, is a satisfactory and consistent way to solve the problem. The expressions are kept relatively simple by writing the basic equations (27) and (31) in terms of the Lagrangian L. However, the full power of the variational principle has not been used in the two-timing analysis and the final results have not yet been justified in general in their most compact form using $\mathscr{L}$.

It is interesting to observe that an exact form of the secular condition can be derived from (32), valid without approximation for small ϵ, and hence valid for all orders. The only argument needed is that R is periodic in θ; therefore, from (32),

$$\frac{\partial}{\partial T}\frac{1}{2\pi}\int_0^{2\pi} P\,d\theta+\frac{\partial}{\partial X}\frac{1}{2\pi}\int_0^{2\pi} Q\,d\theta = 0. \tag{44}$$

Substituting for P and Q from (31), we have

$$\frac{\partial}{\partial T}\frac{1}{2\pi}\int_0^{2\pi} U_\theta L_1(\nu U_\theta+\epsilon U_T,\, \kappa U_\theta+\epsilon U_X,\, U)\,d\theta$$
$$+\frac{\partial}{\partial X}\frac{1}{2\pi}\int_0^{2\pi} U_\theta L_2(\nu U_\theta+\epsilon U_T,\, \kappa U_\theta+\epsilon U_X,\, U)\,d\theta = 0. \tag{45}$$

The successive terms in the expansion of this for $U = \Sigma\epsilon^n U_n$ give the required secular terms in the successive determinations of the U_n. Equation (37) is the lowest-order term in the expansion of (45).

4. Averaging and two-timing a system of conservation equations

In a first attack on slowly varying non-linear waves (Whitham 1965*a*), the equations for (ν, κ, A), and similar overall parameters that arise for higher-order systems, were obtained by averaging an equivalent system of conservation equations. The equivalent system is denoted by

$$\frac{\partial P_i}{\partial t}+\frac{\partial Q_i}{\partial x} = 0 \quad (i = 1, \ldots, m). \tag{46}$$

The intuitive argument is to use their averaged form

$$\frac{\partial}{\partial t}\left(\frac{1}{2\pi}\int_0^{2\pi} P_i\,d\theta\right)+\frac{\partial}{\partial x}\left(\frac{1}{2\pi}\int_0^{2\pi} Q_i\,d\theta\right) = 0, \tag{47}$$

where the averages are computed from the periodic solution.

From the two-timing formalism, (46) becomes

$$\left(\nu\frac{\partial}{\partial\theta}+\epsilon\frac{\partial}{\partial T}\right)P_i+\left(\kappa\frac{\partial}{\partial\theta}+\epsilon\frac{\partial}{\partial X}\right)Q_i = 0,$$

i.e.
$$\frac{\partial R_i}{\partial\theta}+\epsilon\frac{\partial P_i}{\partial T}+\epsilon\frac{\partial Q_i}{\partial X} = 0, \quad R_i = \nu P_i+\kappa Q_i. \tag{48}$$

For each of these we then have

$$R_i^{(0)} = \nu P_i^{(0)}+\kappa Q_i^{(0)} = A_i(X, T), \tag{49}$$

and
$$\frac{\partial}{\partial T}\frac{1}{2\pi}\int_0^{2\pi} P_i^{(0)}\,d\theta+\frac{\partial}{\partial X}\frac{1}{2\pi}\int_0^{2\pi} Q_i^{(0)}\,d\theta = 0. \tag{50}$$

These correspond to (35) and (36), respectively. The integrals in (49) determine the periodic solution with the parameters ν, κ, A_i extended to allow dependence on (X, T). The secular conditions (50) complete the system of equations for ν, κ, A_i. This provides the formal justification for the more intuitive averaging in (47). Again it is observed from (48) that

$$\frac{\partial}{\partial T}\frac{1}{2\pi}\int_0^{2\pi} P_i d\theta + \frac{\partial}{\partial X}\frac{1}{2\pi}\int_0^{2\pi} Q_i d\theta = 0$$

is an exact secular condition.

The method was superseded by the more concise variational approach. This method based on conservation equations is only general if one can be sure that the requisite number of conservation equations (46) always exist. On specific problems, it was always found to be the case, and the experience was sufficiently varied to form the belief that this would always be the case. A direct proof can probably be found, but it is not an easy question. The next section which obtains the equations for ν, κ, A_i without this step implies an indirect verification.

5. Two-timing the variational principle

Most of the discussion will be given for a single function u as in (1), and for one space dimension x. The extensions will then be straightforward.

When u is written in the form (21), (22), the Euler equation becomes

$$\frac{\partial}{\partial \theta}(\nu L_1 + \kappa L_2) - L_3 + \epsilon\frac{\partial L_1}{\partial T} + \epsilon\frac{\partial L_2}{\partial X} = 0, \tag{51}$$

with

$$L_j = L_j(\nu U_\theta + \epsilon U_T, \kappa U_\theta + \epsilon U_X, U), \tag{52}$$

as noted in (27) and (28). It is surprising that this is just the Euler equation of the *three*-variable variational principle

$$\delta \int\int\int_0^{2\pi} L(\nu U_\theta + \epsilon U_T, \kappa U_\theta + \epsilon U_X, U)\, dT\, dX\, d\theta = 0, \tag{53}$$

for the *three*-variable function $U(\theta, X, T)$. In (53) the function U and its variations are taken to be periodic in θ, and the variations in U vanish on the boundary of the (X, T) region. It is even more surprising that (53) is already the averaged variational principle and it is *exact*! What started out as an intuitive argument for the first term of an approximate expansion is not only justified but turns out to contain the whole expansion. There is in fact no assumption that ϵ is small; the only step was to express $u(x, t)$ in the form (21), (22).

The exact averaged Lagrangian is defined by

$$\bar{L} = \frac{1}{2\pi}\int_0^{2\pi} L(\nu U_\theta + \epsilon U_T, \kappa U_\theta + \epsilon U_X, U)\, d\theta, \tag{54}$$

and (53) may be written

$$\delta \int\int \bar{L}\, dT\, dX = 0. \tag{55}$$

Variation of U in (53) gives the whole equation (51). But the simplicity of the arguments in §2 depends on obtaining the equations for the slowly varying parameters by varying those parameters in the average variational principle. Accordingly, we consider variations of (53) with respect to the function $\Theta(X, T)$, which appears in (53) through $\nu = \Theta_T$, $\kappa = \Theta_X$. We have, first,

$$\int\int \{\bar{L}_\nu \delta\Theta_T + \bar{L}_\kappa \delta\Theta_X\} dX\, dT = 0,$$

using (54) for the inner integrals with respect to θ. By the usual variational argument we deduce

$$\frac{\partial}{\partial T}\bar{L}_\nu + \frac{\partial}{\partial X}\bar{L}_\kappa = 0, \tag{56}$$

for arbitrary $\delta\Theta$ which vanish on the boundary of the (X, T) region. This is just the secular condition (45), and it is also exact!

We are now free to use (53) for independent variations of $U(\theta, X, T)$ and $\Theta(X, T)$ to obtain both the equations for U and the secular conditions.

To lowest order, we have

$$\delta \int\int\int_0^{2\pi} L(\nu U_{0\theta}, \kappa U_{0\theta}, U_0)\, dT\, dX\, d\theta = 0. \tag{57}$$

The variation with respect to U_0 gives

$$\frac{\partial}{\partial\theta}(\nu L_1^{(0)} + \kappa L_2^{(0)}) - L_3^{(0)} = 0 \tag{58}$$

(using the notation of (30)) and the variation with respect to Θ gives

$$\frac{\partial}{\partial T}\overline{L_\nu^{(0)}} + \frac{\partial}{\partial X}\overline{L_\kappa^{(0)}} = 0. \tag{59}$$

These are the results (29) and (39) of §3 obtained very simply. Essentially this justifies the theory of §2; it remains to discuss the various forms of the averaged Lagrangian. This is completed in §7, after further amplification of the theory thus far.

The conservation equation (31), and its lowest approximation used in obtaining the integral (35), follow naturally from (53) and (57). In each case the variable θ does not appear explicitly in L. Therefore, the variational principle is invariant if an arbitrary constant is added to θ. Noether's theorem proves that there is a conservation equation corresponding to this invariance; it is

$$\frac{\partial}{\partial\theta}(U_\theta L_{U\theta} - L) + \frac{\partial}{\partial T}(U_\theta L_{UT}) + \frac{\partial}{\partial X}(U_\theta L_{UX}) = 0. \tag{60}$$

With the arguments of L given in (53) it is easily seen that (31) follows. The arguments in (57) give the lowest-order approximation (35). We no longer need (31) for the derivation of the secular condition, but (35) is useful in connexion with (58).

Extension to more variables

The result extends to a system with dependent variables $u_i(\mathbf{x}, t)$ $(i = 1, \ldots, m)$, and with more space dimensions. If the system can be derived from the variational

principle
$$\delta \iint L(u_{it}, u_{i\mathbf{x}}, u_i)\, d\mathbf{x}\, dt = 0, \tag{61}$$
the Euler equations are
$$\frac{\partial}{\partial t} L_{i1} + \frac{\partial}{\partial \mathbf{x}} L_{i2} - L_{i3} = 0 \quad (i = 1, \ldots, m)$$
where
$$L_{i1} = \partial L/\partial u_{it}, \quad L_{i2} = \partial L/\partial u_{i\mathbf{x}}, \quad L_{i3} = \partial L/\partial u_i.$$
After two-timing all the u's in the form
$$u_i(\mathbf{x}, t) = U_i(\theta, \mathbf{X}, T), \quad \theta = \epsilon^{-1}\Theta(\mathbf{X}, T),$$
the Euler equations become
$$\frac{\partial}{\partial \theta}(\nu L_{i1} + \kappa L_{i2}) - L_{i3} + \epsilon \frac{\partial}{\partial T} L_{i1} + \epsilon \frac{\partial}{\partial \mathbf{X}} L_{i2} = 0,$$
where the arguments of the L_{i1}, L_{i2}, L_{i3} are
$$u_{it} = \nu U_{i\theta} + \epsilon U_{iT}, \quad u_{i\mathbf{x}} = \boldsymbol{\kappa} U_{i\theta} + \epsilon U_{i\mathbf{X}}, \quad u_i = U_i.$$
These are the Euler equations for
$$\delta \iiint_0^{2\pi} L(\nu U_{i\theta} + \epsilon U_{iT}, \boldsymbol{\kappa} U_{i\theta} + \epsilon U_{i\mathbf{X}}, U_i)\, d\mathbf{X}\, dT\, d\theta = 0. \tag{62}$$

In applications of these methods to such higher-order systems, it was noted in earlier papers that some of the u_i appear only through their derivatives in (61). Typically, such u_i are potentials, whose derivatives are physical quantities. Let one of these be denoted by ϕ. In the uniform periodic wave train, for maximum generality ϕ must be expressed as
$$\phi = \beta x + \gamma t + \Phi(\theta), \quad \theta = \kappa x + \nu t, \tag{63}$$
where $\Phi(\theta)$ is periodic and β, γ are arbitrary constants. The derivatives of ϕ, which are the physical quantities, are then periodic functions of θ. The pair (β, γ) behave like (κ, ν) in the analysis. They have been called pseudo wave-numbers and frequencies, since they give the increments in ϕ over one period or wavelength, just as (κ, ν) give the increments of θ.

In the slowly varying wave train, (63) is generalized to
$$\left.\begin{aligned} &\phi = \psi + \Phi(\theta, X, T), \\ &\theta = \epsilon^{-1}\Theta(X, T), \quad \psi = \epsilon^{-1}\Psi(X, T), \\ &\nu = \Theta_T, \quad \kappa = \Theta_X, \quad \gamma = \Psi_T, \quad \beta = \Psi_X. \end{aligned}\right\} \tag{64}$$
The average Lagrangian in (62) now includes dependence on $\Psi(X, T)$ through $\gamma = \Psi_T$, $\beta = \Psi_X$. The variations with respect to Ψ give additional secular conditions
$$\frac{\partial}{\partial T} \bar{L}_\gamma + \frac{\partial}{\partial X} \bar{L}_\beta = 0, \tag{65}$$
and these are required to complete the solution to lowest order. The details are not given here in general; typical cases can be seen in the previous papers (Whitham 1965, 1967) and will be included in the examples of §8.

6. Relation between two-timing and averaging

It is intriguing to investigate further the relation of the precise result (53) with the intuitive ideas of averaging. One might also question whether (53) can be posed directly without any reference to the Euler equations, even though their use in § 5 has been held to a minimum.

With the choice of the functional form (21) and (22) the original variational principle (1) becomes

$$\delta \int\int [L(\nu U_\theta + \epsilon U_T, \kappa U_\theta + \epsilon U_X, U)]_{\theta = \epsilon^{-1}\Theta} dX\, dT = 0, \tag{66}$$

where $[\ ]_{\theta=\epsilon^{-1}\Theta}$ denotes that, at this first step, θ is still linked to (X, T) through the substitution $\theta = \epsilon^{-1}\Theta(X, T)$. The idea of the averaging process is that the dependence on ϵ through $\theta = \epsilon^{-1}\Theta$ is highly oscillatory so that the integrand in (66) may be replaced by its mean value. A consistent way to introduce this is to use Fourier series. For functions $U(\theta, X, T, \epsilon)$ periodic in θ,

$$L(\nu U_\theta + \epsilon U_T, \kappa U_\theta + \epsilon U_X, U)$$

is also periodic in θ and we may introduce its Fourier series

$$L(\nu U_\theta + \epsilon U_T, \kappa U_\theta + \epsilon U_X, U) = C_0 + \sum_{n=1}^{\infty} (C_n \cos n\theta + S_n \sin n\theta). \tag{67}$$

The coefficients C_n, S_n are functionals of $U(\theta, X, T, \epsilon)$, and of $\Theta(X, T)$ through $\nu = \Theta_T$, $\kappa = \Theta_X$. In detail

$$C_0 = \frac{1}{2\pi}\int_0^{2\pi} L(\nu U_\theta + \epsilon U_T, \kappa U_\theta + \epsilon U_X, U)\, d\theta, \tag{68}$$

and

$$C_n = \frac{1}{\pi}\int_0^{2\pi} L \cos n\theta\, d\theta, \quad S_n = \frac{1}{\pi}\int_0^{2\pi} L \sin n\theta\, d\theta, \tag{69}$$

(with the same arguments for L). Then, (66) becomes

$$\delta \int\int \left\{ C_0 + \sum_{n=1}^{\infty} \left(C_n \cos\frac{n\Theta}{\epsilon} + S_n \sin\frac{n\Theta}{\epsilon} \right) \right\} dX\, dT = 0. \tag{70}$$

Comparing (53) and (68), we see that (53) is in fact

$$\delta \int\int C_0\, dX\, dT = 0. \tag{71}$$

Since (71) is therefore exact, it must be valid to omit the other Fourier terms in (70).

The basis of the intuitive argument is that, for small ϵ, the oscillatory terms in the integral in (70) contribute very little and may be omitted. Indeed, by repeated integration by parts they may be made of smaller order than *any* power of ϵ. If they are then neglected in (70) before taking the variations, we have (71) of course. However, if they are retained, they now involve high derivatives in Θ and U, and the usual variational argument to establish the Euler equations

will undo the integration by parts and we are back to (70). Moreover, variations of Θ in (70) will include terms of order ϵ^{-1}, from the variations of $\cos n\epsilon^{-1}\Theta$ and $\sin n\epsilon^{-1}\Theta$, and these would appear to be the dominant terms. Yet we know already that (71) is correct.

There is then an apparent lack of uniqueness with respect to the ordering of the terms in (70) with respect to ϵ. The only way out of this disturbing situation would be if each of the Fourier terms in (70) is individually stationary, i.e.

$$\delta \iint \left(C_n \cos\frac{n\Theta}{\epsilon} + S_n \sin\frac{n\Theta}{\epsilon}\right) dX\,dT = 0, \tag{72}$$

for all n. Surprisingly enough this turns out to be true.

To prove it, we show that the Euler equations for (72), as well as those for (71), are equivalent to (51). The coefficients C_n, S_n are defined by (69). Accordingly (72) may be written

$$\delta \iint \int_0^{2\pi} \cos n\left(\theta - \frac{\Theta}{\epsilon}\right) L(\nu U_\theta + \epsilon U_T,\, \kappa U_\theta + \epsilon U_X,\, U)\,dX\,dT\,d\theta = 0. \tag{73}$$

The Euler equation for variations in U is

$$\frac{\partial}{\partial\theta}\left\{(\nu L_1 + \kappa L_2)\cos n\left(\theta - \frac{\Theta}{\epsilon}\right)\right\} - L_3 \cos n\left(\theta - \frac{\Theta}{\epsilon}\right)$$
$$+ \epsilon\frac{\partial}{\partial T}\left\{L_1 \cos n\left(\theta - \frac{\Theta}{\epsilon}\right)\right\} + \epsilon\frac{\partial}{\partial X}\left\{L_2 \cos n\left(\theta - \frac{\Theta}{\epsilon}\right)\right\} = 0. \tag{74}$$

This equation may be expanded to

$$\left\{\frac{\partial}{\partial\theta}(\nu L_1 + \kappa L_2) - L_3 + \epsilon\frac{\partial L_1}{\partial T} + \epsilon\frac{\partial L_2}{\partial X}\right\}\cos n\left(\theta - \frac{\Theta}{\epsilon}\right) = 0. \tag{75}$$

The expression in brackets is just (51) and the result is proved.

Again all the information is in the equation for U, but we may consider the variations of (73) with respect to Θ. The result is

$$\frac{\partial}{\partial T}\frac{1}{2\pi}\int_0^{2\pi} U_\theta L_1 \cos n\left(\theta - \frac{\Theta}{\epsilon}\right) d\theta + \frac{\partial}{\partial X}\int_0^{2\pi} U_\theta L_2 \cos n\left(\theta - \frac{\Theta}{\epsilon}\right) d\theta$$
$$- \frac{1}{2\pi}\int_0^{2\pi} \frac{nL}{\epsilon}\sin n\left(\theta - \frac{\Theta}{\epsilon}\right) d\theta = 0.$$

It may be expanded to read

$$\left\{\epsilon\frac{\partial P_n^c}{\partial T} + \epsilon\frac{\partial Q_n^c}{\partial X} + nR_n^s\right\}\cos\frac{n\Theta}{\epsilon} + \left\{\epsilon\frac{\partial P_n^s}{\partial T} + \epsilon\frac{\partial Q_n^s}{\partial X} - nR_n^c\right\}\sin\frac{n\Theta}{\epsilon} = 0, \tag{76}$$

where P, Q, R refer to the quantities introduced in (31), (32) and superscripts P_n^c, P_n^s, etc., denote the cosine and sine coefficients in their Fourier series. It is easily shown from (32), by a continuation of the secular argument, that the bracketed terms in (76) vanish separately. For the secular condition is that P, Q, R must be periodic in θ. Hence, these quantities can be expanded in their Fourier series

$$P = P_0 + \sum_{n=1}^{\infty} (P_n^c \cos n\theta + P_n^s \sin n\theta), \text{ etc.} \tag{77}$$

Substitution of these in (32) gives

$$\left.\begin{aligned} \frac{\partial P_0}{\partial T}+\frac{\partial Q_0}{\partial X} &= 0, \\ \epsilon\frac{\partial P_n^c}{\partial T}+\epsilon\frac{\partial Q_n^c}{\partial X} &= -nR_n^s, \\ \epsilon\frac{\partial P_n^s}{\partial T}+\epsilon\frac{\partial Q_n^s}{\partial X} &= nR_n^c. \end{aligned}\right\} \tag{78}$$

The first of these is the main secular condition noted in (44); the additional ones add nothing new in the lowest order.

7. Optimum form of the average Lagrangian

The variational principle (57) and its consequences in (58) and (59) have been justified. We now consider the various manipulations in order to obtain the most effective form. The discussion throughout this section will refer to the lowest-order approximation U_0 and the subscript zero will be dropped; the corresponding superscript zero will be dropped in (58) and (59).

In the first approximation, then, we have

$$\delta\int\int \bar{L}\,dX\,dT = 0, \tag{79}$$

where

$$\bar{L} = \frac{1}{2\pi}\int_0^{2\pi} L(\nu U_\theta, \kappa U_\theta, U)\,d\theta, \tag{80}$$

$$\nu = \Theta_T, \quad \kappa = \Theta_X.$$

The variation with respect to $U(\theta, X, T)$ gives

$$\frac{\partial}{\partial\theta}(\nu L_1+\kappa L_2)-L_3 = 0; \tag{81}$$

the variation with respect to $\Theta(X, T)$ gives

$$\frac{\partial}{\partial T}\bar{L}_\nu+\frac{\partial}{\partial X}\bar{L}_\kappa = 0. \tag{82}$$

Equation (81) has the integral

$$(\nu L_1+\kappa L_2)U_\theta-L = A(X, T). \tag{83}$$

The aim is to use this integral to evaluate $\bar{L}$ as a function of (ν, κ, A).

The method is essentially a Hamiltonian version of the equations. The quantity U_θ is eliminated in favour of $\partial L/\partial U_\theta$ just as $\dot{q}$ is eliminated in favour of a generalized momentum $p = \partial L/\partial\dot{q}$ in ordinary dynamics. A new variable Π is defined by

$$\Pi = \partial L/\partial U_\theta = \nu L_1+\kappa L_2, \tag{84}$$

and the 'Hamiltonian' $H(\Pi, U; \Theta)$ is defined by

$$H = U_\theta\partial L/\partial U_\theta-L = U_\theta(\nu L_1+\kappa L_2)-L. \tag{85}$$

It is noticed immediately that (83) is the 'energy integral'

$$H(\Pi, U; \Theta) = A(X, T). \tag{86}$$

The Hamiltonian is a function of Π and U, but it is a functional of Θ since it depends on $\nu = \Theta_T$ and $\kappa = \Theta_X$. The semi-colon notation will be used in displaying the arguments to make this distinction.

From the transformation defined by (84) and (85),

$$U_\theta = \partial H/\partial \Pi, \tag{87}$$

and (81) becomes

$$\Pi_\theta = -\partial H/\partial U. \tag{88}$$

These are the 'Hamiltonian equations' to replace 'Lagrange's equation' (81). The variational principle (79) may now be written with

$$\bar{L} = \frac{1}{2\pi}\int_0^{2\pi} (\Pi U_\theta - H)\, d\theta. \tag{89}$$

The *independent* variations of Π and U lead to the equations (87) and (88), respectively. The independence of the variations of Π and U is crucially important. It is an extension of the original form, because in (79), (80) the variation δU_θ, and hence $\delta\Pi$, is determined in terms of δU. Yet, we see that (87) and (88) do follow from (89), with $\delta\Pi$ and δU independent. Thus, we may use the extended form. It is exactly the extension used in ordinary mechanics.

We also saw that the stationary values of Π and U satisfy the integral (86). We may, therefore, take the stationary value for (89) from the class of functions Π, U that satisfy (86). It is no longer too restrictive to use (86) in its entirety, because the dispersion relation cannot be inferred without also using the relation of Π to U_θ. The latter is given by (87) and we do not use that. It should be emphasized again that this is only possible by the extension that Π and U can be varied independently in (89).

Finally, then, we solve (86) in the form

$$\Pi = \Pi(U, A; \Theta)$$

and (89) may be written

$$\bar{L} = \mathscr{L}\{A; \Theta\} = \frac{1}{2\pi}\oint \Pi\, dU - A. \tag{90}$$

$\mathscr{L}\{A;\, \Theta\}$ is a functional of Θ and a function of A; it is a function $\mathscr{L}(\nu, \kappa, A)$ of $\nu = \Theta_T$, $\kappa = \Theta_X$ and A.

Extension to higher-order systems

The generalization to cases with more functions u_i is carried out similarly. The exact averaged Lagrangian is given in (62). As noted in (64) some of the u_i may be potentials and require that two-timing form. It will suffice to explain the situation for higher-order systems in the case of two functions u_i, one of which is a potential. They will be denoted by u and ϕ. The Lagrangian is

$$L = L(u_t, u_x, u, \phi_t, \phi_x)$$

with no explicit dependence on ϕ. To lowest order the variational principle is (79) with

$$\bar{L} = \frac{1}{2\pi}\int_0^{2\pi} L(\nu U_\theta, \kappa U_\theta, U, \gamma+\nu\Phi_\theta, \beta+\kappa\Phi_\theta)\,d\theta, \tag{91}$$

$$\nu = \Theta_T, \quad \kappa = \Theta_X, \quad \gamma = \Psi_T, \quad \beta = \Psi_X. \tag{92}$$

In addition to (81), we have the Euler equation from variations in Φ:

$$\frac{\partial}{\partial\theta}(\nu L_4+\kappa L_5) = 0. \tag{93}$$

The absence of explicit dependence on Φ in (91) is significant in that (93) integrates immediately, and provides a second integral to use with (83). The variations of Ψ give a second secular condition

$$\frac{\partial}{\partial T}\bar{L}_\gamma+\frac{\partial}{\partial X}\bar{L}_\beta = 0 \tag{94}$$

to add to (82).

The Hamiltonian form is obtained from the transformation

$$\Pi_1 = \frac{\partial L}{\partial U_\theta} = \nu L_1+\kappa L_2, \tag{95}$$

$$\Pi_2 = \frac{\partial L}{\partial \Phi_\theta} = \nu L_4+\kappa L_5, \tag{96}$$

$$H(\Pi_1, \Pi_2, U, \Phi; \Theta, \Psi) = \Pi_1 U_\theta+\Pi_2\Phi_\theta-L. \tag{97}$$

The Euler equations become

$$U_\theta = \frac{\partial H}{\partial \Pi_1}, \quad \Phi_\theta = \frac{\partial H}{\partial \Pi_2}, \quad \Pi_{1\theta} = -\frac{\partial H}{\partial U}, \quad \Pi_{2\theta} = -\frac{\partial H}{\partial \Phi}, \tag{98}$$

and follow from (91) with

$$\bar{L} = \frac{1}{2\pi}\int_0^{2\pi}(\Pi_1 U_\theta+\Pi_2\Phi_\theta-H)\,d\theta. \tag{99}$$

Again a crucial extension is that Π_1, Π_2, U, Φ may be varied independently.

Since Φ was in fact absent from L, it is also absent from H and the last equation in (98) gives

$$\Pi_2 = B(X, T). \tag{100}$$

This is just the integral obtained from (93). The other one is

$$H = A(X, T),$$

which (with (100)) can be solved to give

$$\Pi_1 = \Pi_1(U, A, B; \Theta, \Psi). \tag{101}$$

In view of (100), the term

$$\int_0^{2\pi}\Pi_2\Phi_\theta\,d\theta = B\int_0^{2\pi}\Phi_\theta\,d\theta = 0 \tag{102}$$

in (99), since Φ is periodic. Hence,

$$\bar{L} = \mathscr{L}\{A, B; \Theta, \Psi\} = \frac{1}{2\pi}\oint \Pi_1\,dU-A, \tag{103}$$

where Π_1 is given by (101). The derivatives of Θ, Ψ appear in $\mathscr{L}$.

The required equations are then

$$\mathscr{L}_A = 0, \quad \mathscr{L}_B = 0, \tag{104}$$

$$\frac{\partial}{\partial T}\mathscr{L}_\nu + \frac{\partial}{\partial X}\mathscr{L}_\kappa = 0, \quad \frac{\partial}{\partial T}\mathscr{L}_\gamma + \frac{\partial}{\partial X}\mathscr{L}_\beta = 0. \tag{105}$$

Since the additional variable ϕ was taken to be a potential, it automatically brought in a second integral to simplify (99). In more general systems, one might expect sums of terms like $\Pi_1 U_\theta$ and $\Pi_2 \Phi_\theta$ in (99). All the Π_2 will be integrals and the terms $\Pi_2 \Phi_\theta$ integrate to zero as in (102). We then have

$$\bar{L} = \frac{1}{2\pi}\int_0^{2\pi} \{\Sigma \Pi_1 U_\theta\} d\theta - A,$$

$$\Pi_2 = B, \quad H = A. \tag{106}$$

Further simplification down to integrals of the type in (90) would depend on finding further integrals. In all known examples, all the u_i except one are potentials. So there is only one U and Π_1; the form in (103) applies. Even if eventually other types arise, the final simplification is not essential; the method still applies.

Linear problems

For linear problems, one need not resort to this transformation to Hamiltonian form. Its object is to introduce some restriction on the *form* of the periodic solution into $\bar{L}$ while leaving enough flexibility to apply the variational arguments. In linear problems, the U_i are sinusoidal in θ. One can introduce the appropriate forms for them without being overly restrictive, provided the dispersion relation is not used. In non-linear problems, the form of the U, the dispersion relation and the amplitudes are coupled together. These require the more ingenious treatment of this section.

8. Illustrative examples

(a) The Klein–Gordon equation

The Lagrangian is $$L = \tfrac{1}{2}u_t^2 - \tfrac{1}{2}u_x^2 - V(u),$$

and the lowest-order approximation is

$$L = \tfrac{1}{2}(\nu^2 - \kappa^2)U_\theta^2 - V(U).$$

The transformation (84), (85) to Hamiltonian form is

$$\Pi = \partial L/\partial U_\theta = (\nu^2 - \kappa^2)U_\theta$$

$$H = U_\theta \partial L/\partial U_\theta - L = \tfrac{1}{2}(\nu^2 - \kappa^2)U_\theta^2 + V(U)$$

$$= \tfrac{1}{2}(\nu^2 - \kappa^2)^{-1}\Pi^2 + V(U).$$

The integral $H = A$ is solved as

$$\Pi = 2^{\frac{1}{2}}(\nu^2 - \kappa^2)^{\frac{1}{2}}[A - V(U)]^{\frac{1}{2}},$$

and
$$\mathscr{L} = \frac{1}{2\pi}\oint \Pi\, dU - A$$
$$= \frac{1}{2\pi}[2(\nu^2-\kappa^2)]^{\frac{1}{2}}\oint [A - V(U)]^{\frac{1}{2}} dU - A,$$
in agreement with (20).

In the linear case, $V(U) = \frac{1}{2}U^2$, the integral in $\mathscr{L}$ can be evaluated explicitly and we have
$$\mathscr{L} = [(\nu^2-\kappa^2)^{\frac{1}{2}} - 1]\,A. \tag{107}$$
From the energy equation, A is proportional to a^2, and this result conforms with the general result in (12).

(*b*) *Linear example*

To illustrate the remarks on linear problems at the end of §7, we do the linear Klein–Gordon case directly. We have
$$\bar{L} = \frac{1}{2\pi}\int_0^{2\pi} \{\tfrac{1}{2}(\nu^2-\kappa^2)U_\theta^2 - \tfrac{1}{2}U^2\}\, d\theta.$$
We substitute
$$U = a\cos\theta,$$
and deduce
$$\bar{L} = \mathscr{L}^* = \tfrac{1}{4}(\nu^2-\kappa^2-1)\,a^2. \tag{108}$$

The dispersion relation is written in a different form in (107) and (108), but the final results are equivalent. In any linear problem, $\mathscr{L} = F(\omega,\kappa)\,a^2$, and the slowly varying equations are
$$\mathscr{L}_a \propto F(\omega,\kappa) = 0, \tag{109}$$
$$\frac{\partial}{\partial T}(F_\omega a^2) + \frac{\partial}{\partial X}(F_\kappa a^2) = 0, \tag{110}$$
$$\frac{\partial \kappa}{\partial T} + \frac{\partial \omega}{\partial X} = 0. \tag{111}$$
We suppose (109) is solved in the form $\omega = W(\kappa)$. Then $F(W,\kappa) = 0$ is an identity and it follows that
$$C(\kappa)\,F_\omega(W,\kappa) + F_\kappa(W,\kappa) = 0,$$
where
$$C(\kappa) = W'(\kappa)$$
is the group velocity. Equations (110), (111) may then be written
$$\frac{\partial}{\partial T}\{a^2 f(\kappa)\} + \frac{\partial}{\partial X}\{a^2 f(\kappa)\,C(\kappa)\} = 0, \tag{112}$$
$$\frac{\partial \kappa}{\partial T} + C(\kappa)\frac{\partial \kappa}{\partial X} = 0, \tag{113}$$
where $f(\kappa) = F_\omega(W,\kappa)$. Equation (112) may be expanded to
$$f(\kappa)\left\{\frac{\partial a^2}{\partial T} + \frac{\partial (Ca^2)}{\partial X}\right\} + f'(\kappa)\left\{\frac{\partial \kappa}{\partial T} + C\frac{\partial \kappa}{\partial X}\right\} = 0.$$
The second term is zero by (113). Finally, therefore, the equations reduce to
$$\frac{\partial a^2}{\partial T} + \frac{\partial}{\partial X}(a^2 C) = 0, \quad \frac{\partial \kappa}{\partial T} + C(\kappa)\frac{\partial \kappa}{\partial X} = 0.$$

This final form shows that different choices of $F(\omega, \kappa)$ will lead to the same final equations provided that they have the same solution $\omega = W(\kappa)$.

(c) *Boussinesq equations for water waves*

In Boussinesq's approximation for long water waves the Lagrangian is

$$L = \tfrac{1}{2}u_t^2 - \tfrac{1}{2}gu^2 - u(\phi_t + \tfrac{1}{2}\phi_x^2),$$

where u is the water height and ϕ is a velocity potential (see Whitham 1965*b*). The notation u for the height is to conform with the general notation in this paper. To lowest order,

$$L = \tfrac{1}{2}\nu^2 U_\theta^2 - \tfrac{1}{2}gU^2 - U\{\gamma + \nu\Phi_\theta + \tfrac{1}{2}(\beta + \kappa\Phi_\theta)^2\}.$$

Introducing the Hamiltonian transformation, as in (95) etc., we have

$$\Pi_1 = \frac{\partial L}{\partial U_\theta} = \nu^2 U_\theta, \quad \Pi_2 = \frac{\partial L}{\partial \Phi_\theta} = -U\{\nu + \kappa(\beta + \kappa\Phi_\theta)\},$$

$$H = \tfrac{1}{2}\nu^{-2}\Pi_1^2 - \tfrac{1}{2}\kappa^{-2}U^{-1}\{\Pi_2 + U(\nu + \beta\kappa)\}^2 + U(\gamma + \tfrac{1}{2}\beta^2) + \tfrac{1}{2}gU^2.$$

Since Φ is absent from H, $\qquad \Pi_2 = B(X, T)$.

Then, solving $\qquad H = A(X, T)$
for Π_1, we have

$$\mathscr{L} = \frac{1}{2\pi}\oint \Pi_1 dU - A.$$

$$= \frac{1}{2\pi}\nu\oint \{\kappa^{-2}U^{-1}(B + U\nu + U\beta\kappa)^2 - U(2\gamma + \beta^2) - gU^2 + 2A\}^{\frac{1}{2}} dU - A.$$

This is the result obtained previously (Whitham 1965*b*) with slight changes in notation.

The variational equations are (104), (105).

(d) *Central orbits*

The methods developed here for waves have required both extensions and new versions of ideas and techniques in mechanics. It is interesting to return to some of the questions in mechanics from this different point of view. A typical problem arises in central orbits and the theory of their adiabatic invariants. The theory of this paper applies with the simplification that the dependent variables u_i are functions of t alone. The slow variations are introduced, now, by slow variations in time of the central force field.

If r and ϕ are the radius and polar angle, respectively, the Lagrangian for an orbit in a force field with potential $V(r, \epsilon t)$ is

$$L = \tfrac{1}{2}\dot{r}^2 + \tfrac{1}{2}r^2\dot{\phi}^2 - V(r, \epsilon t).$$

The Euler equations are $\qquad \ddot{r} - r\dot{\phi}^2 + V_r(r, \epsilon t) = 0,$

$$\frac{d}{dt}(r^2\dot{\phi}) = 0.$$

For $\epsilon = 0$, the solution may be written

$$r = R(\theta), \quad \theta = \nu t,$$
$$\phi = \gamma t + \Phi(\theta),$$

where R and Φ are periodic functions of θ with period 2π. In this unperturbed case, ν and γ are constant. The period of r in time is $2\pi/\nu$. In that period, ϕ increases by $2\pi\gamma/\nu$; the orbit is closed if $\gamma = \nu$.

For the perturbed case $\epsilon \neq 0$, the solution is generalized to the slowly varying form

$$r = R(\theta, T), \quad T = \epsilon t, \quad \theta = \epsilon^{-1}\Theta(T), \quad \nu = \Theta_T$$
$$\phi = \psi + \Phi(\theta, T), \quad \psi = \epsilon^{-1}\Psi(T), \quad \gamma = \Psi_T.$$

To lowest order the Lagrangian is

$$L = \tfrac{1}{2}\nu^2 R_\theta^2 + \tfrac{1}{2}R^2(\gamma + \nu\Phi_\theta)^2 - V(R, T).$$

The 'Hamiltonian transformation' is

$$\Pi_1 = \partial L/\partial R_\theta = \nu^2 R_\theta, \quad \Pi_2 = \partial L/\partial \Phi_\theta = \nu R^2(\gamma + \nu\Phi_\theta),$$
$$H = \tfrac{1}{2}\nu^{-2}\Pi_1^2 + \tfrac{1}{2}\nu^{-2}R^{-2}\Pi_2^2 - \nu^{-1}\gamma\Pi_2 + V(R, T).$$

(This 'Hamiltonian' differs by the term $\nu^{-1}\gamma\Pi_2$ from the usual one.) As before, Π_2 and H are integrals as regards the θ dependence. It is convenient to introduce corresponding parameters $M(T)$, $E(T)$, by

$$\Pi_2 = \nu M,$$
$$H = E - \gamma M,$$

so that M and E are the angular momentum and energy, respectively. These are solved for Π_1 and Π_2 to give

$$\Pi_1 = \nu\{2E - M^2R^{-2} - 2V(R, T)\}^{\frac{1}{2}}, \quad \Pi_2 = \nu M.$$

The average Lagrangian is

$$\bar{L} = \frac{1}{2\pi}\int_0^{2\pi} \{\Pi_1 R_\theta + \Pi_2\Phi_\theta - H\}\,d\theta.$$

Therefore $$\mathscr{L} = \frac{\nu}{2\pi}\oint \{2E - M^2R^{-2} - 2V(R, T)\}^{\frac{1}{2}}\,dR + \gamma M - E.$$

The variations with respect to Θ and Ψ give

$$\frac{d}{dT}\mathscr{L}_\nu = 0, \quad \frac{d}{dT}\mathscr{L}_\gamma = 0.$$

Hence, $$I_1 = \mathscr{L}_\nu = \frac{1}{2\pi}\oint \{2E - M^2R^{-2} - 2V(R, T)\}^{\frac{1}{2}}\,dR,$$
$$I_2 = \mathscr{L}_\gamma = M,$$

are the adiabatic invariants which remain constant for slow changes in the force field. The second shows that the angular momentum remains constant, as it must in a central force field. The first then determines the slow changes in the energy E.

The variations of $\mathscr{L}$ with respect to E and M give

$$\mathscr{L}_E = \nu\partial I_1/\partial E - 1 = 0, \quad \mathscr{L}_M = \nu\partial I_1/\partial M + \gamma = 0.$$

These determine the frequencies ν and γ.

A convenient alternative form is to introduce I_1 and I_2 as parameters in place of E and M. Then

$$\mathscr{L} = \nu I_1 + \gamma I_2 - E(I_1, I_2),$$

and the variational equations are

$$\frac{dI_1}{dT} = 0, \quad \frac{dI_2}{dT} = 0,$$

$$\nu = \frac{\partial E}{\partial I_1}, \quad \gamma = \frac{\partial E}{\partial I_2}.$$

These are the standard results of the theory of adiabatic invariants, but they are usually obtained by other methods. A good account is given by Landau & Lifshitz (1960).

For the case of an inverse square force, the potential energy is $V = -\alpha(T)/R$. For this case the integral in $\mathscr{L}$ can be evaluated explicitly (most simply by contour integration) and we have

$$\mathscr{L} = \nu\left(\frac{\alpha}{[2|E|]^{\frac{1}{2}}} - M\right) + \gamma M - E.$$

The adiabatic invariants are

$$I_1 = \frac{\alpha}{[2|E|]^{\frac{1}{2}}} - M, \quad I_2 = M.$$

The frequency equations $\mathscr{L}_E = \mathscr{L}_M = 0$ give

$$\nu = \gamma = (-2E)^{\frac{3}{2}}/\alpha.$$

The energy may be expressed in terms of I_1, I_2 by

$$E = -\frac{\alpha^2}{2(I_1 + I_2)^2}.$$

Since the frequencies are equal, the orbit is closed. As regards the general theory, the equal frequencies make this a degenerate case.

This research was supported by the Office of Naval Research, U.S. Navy.

REFERENCES

COLE, J. D. 1968 *Perturbation Methods in Applied Mathematics*. Waltham, Mass.: Blaisdell.

KUZMAK, G. E. 1959 *Prikl. Mat. Mekh. Akad. Nauk S.S.S.R.* **23**, 515. (Translated in *Appl. Math. Mech.* **23**, 730.)

LANDAU, L. D. & LIFSHITZ, E. M. 1960 *Mechanics*. Oxford: Pergamon.

LUKE, J. C. 1966 *Proc. Roy. Soc.* A **292**, 403.

WHITHAM, G. B. 1965*a* *Proc. Roy. Soc.* A **283**, 238.

WHITHAM, G. B. 1965*b* *J. Fluid Mech.* **22**, 273.

WHITHAM, G. B. 1967*a* *J. Fluid Mech.* **27**, 399.

WHITHAM, G. B. 1967*b* *Proc Roy. Soc.* A **299**, 6.