J. Fluid Mech. (1986), *vol.* 170, *pp.* 21–51
Printed in Great Britain

Amplitude propagation in slowly varying trains of shear-flow instability waves

By J. M. RUSSELL

School of Aerospace, Mechanical, and Nuclear Engineering, The University of Oklahoma, Norman, OK 73019, USA

(Received 5 March 1985 and in revised form 10 February 1986)

The analog of Whitham's law of conservation of wave action density is derived in the case of Rayleigh instability waves. The analysis allows for wave propagation in two space dimensions, non-unidirectionality of the background flow velocity profiles and weak horizontal nonuniformity and unsteadiness of those profiles. The small disturbance equations of motion in the Eulerian flow description are subject to a change of dependent variable in which the new variable represents the pressure-driven part of a disturbance material coordinate function as a function of the Cartesian spatial coordinates and time. Several variational principles expressing the physics of the small disturbance equations of motion are presented in terms of this new variable. A law of conservation of 'bilinear wave action density' is derived by a method intermediate between those of Jimenez and Whitham (1976) and Hayes (1970*a*). The distinction between the observed square amplitude of an amplified wavetrain and the wave action density is discussed. Three types of algebraic focusing are discussed, the first being the far-field 'caustics', the second being near-field 'movable singularities', and the third being a focusing mechanism due to Landahl (1972) which we here derive under somewhat weaker hypotheses.

1. Introduction

The present work is a contribution to the theory of inviscid shear-flow instability waves, particularly those whose dynamics are governed locally by the Rayleigh stability equation. Our main concern will be the effects of weak non-uniformity and unsteadiness of the parameters of the wavetrain (such as frequency and wavenumber) on the evolution of the local square amplitude of a train or packet of such waves. A comprehensive presentation of the theory of Rayleigh waves may be found in the article by Drazin & Howard (1966). Most of the results derived there apply to constant-parameter wavetrains, the formulation of the initial-value problem by a Laplace transform method being a notable exception.

Employing the nomenclature of Hayes (1970*a*), we regard the set of independent variables in the set of partial differential equations describing some physical problem as the *physical space*. The physical space is the product of a *propagation space* and a *cross space*, where the propagation space is the set of independent variables upon which solutions may exhibit the kind of oscillatory behaviour associated with wave motion. If the propagation space of a system coincides with its physical space, then the waves are called *local* waves. Otherwise, they are called *modal* waves. Shear-flow instability waves, since they do not propagate in the direction transverse to the undisturbed flow, are modal waves under this classification.

The theory of weakly non-uniform linear dispersive waves presented in chapter 11 of Whitham (1974) is concerned, for the most part, with conservative systems of local waves, where for present purposes we regard a conservative system as one whose dispersion relation (relating frequency to wavenumber) involves only real quantities. In the absence of modifications, therefore, Whitham's 1974 theory is not applicable to shear-flow instability waves which are modal and nonconservative.

Whitham's variational method for the analysis of amplitude propagation is extremely attractive, however, and suggests that a modification to the theory to allow for non-conservative modal waves would be of value. Hayes (1970*a*) has presented an alternative variational method for analysis of amplitude propagation which allows for waves of modal type, though the restriction to systems governed by a real dispersion relation is retained.

Local waves in linear problems will be conservative if the orders of the partial derivatives in the partial differential equations of motion are all even or all odd (if a term in the equation of motion involves a mixed partial derivative, we regard the total order of that term as the relevant one for present purposes). Conversely, local waves will be non-conservative only if there exists at least one 'dissipative' term which is identified by its non-conformity to the 'all even or all odd' rule. Jimenez & Whitham (1976) have presented a modified version of Whitham's variational method that allows for dissipative terms of this type.

Rayleigh waves, by contrast, are non-conservative for reasons which have no direct relation to any explicit dissipative terms in the equations of motion. They are modal waves whose equation of motion in the cross-space variable (i.e. the Rayleigh stability equation) exhibits regular singular points of logarithmic type (i.e. 'critical layers'). If the curvature of the mean velocity profile is non-zero at a critical layer, at least one homogeneous solution (say ϕ_n) must exist whose derivative has discontinuous imaginary part across it. The typical manner in which such increments in $\mathrm{Im}\{\phi_n'(y)\}$ are reconciled with homogeneous boundary conditions is to admit appropriate complex parameters in the Rayleigh equation. Thus, it is the complex-valued increments of the homogeneous solutions rather than any explicit dissipative terms in the equations of motion that accounts for complex numbers in the dispersion relation.

Extensions of Whitham's method to allow for non-conservative wavetrains whose non-conservative character is due to logarithmic singularities in the differential equations of motion in the cross-space have not appeared in the literature. The present contribution, which incorporates many of the ideas of Hayes (1970*a*) theory of conservative modal waves with Jimenez & Whitham's (1976) theory of non-conservative local ones, is meant to supply such an extension.

At least two effective methods of analysis have been applied by previous investigators to the problem of amplitude propagation in shear-flow instability waves. In one approach (cf. Chin 1980; Landahl 1982), the complex dispersion relation expressing frequency as a function of wavenumber in the stability problem is taken as the starting point. By appeal to the Fourier integral theorem, the relationship between an initial wavenumber distribution function and the dependent variable ψ in an associated physical problem may be written down (as is done, for example, in classical applications of the method of stationary phase, cf. Lighthill 1978 §3.7). Expanding the dispersion relation in a complete Taylor series in the wavenumber, an infinite-term partial differential equation for ψ may be derived. Chin (1980) applied the WKBJ method to the solution of the equation for ψ and succeeded in analytically summing (to a few orders in the WKBJ expansion) the infinite series associated with the original power-series expansion in the wavenumber. Among the results derived by Chin was an equation for the propagation of the local square amplitude of a slowly varying wavetrain.

The second main approach to amplitude propagation in shear waves is the one represented by the work of Nayfeh (1980) and Itoh (1980, 1981), which we will call 'the direct asymptotic approach'. In this approach, no use is made of the Fourier integral theorem. Solutions of the small-disturbance equations of motion are sought directly in the form of slowly varying wavetrains by either a WKBJ approximation (in Itoh's approach) or by a multiple scale expansion (in Nayfeh's). In each case, the lowest-order balance in the equations of motion leads to the homogeneous two-point boundary-value problem (involving derivatives with respect to one cross-space variable) that would normally be encountered if non-uniformity of the parameters of the wavetrain were ignored. The next higher-order balance involves the solution of an inhomogeneous equation (or a set of equations) whose homogeneous operator is the same as that of the lowest-order problem. Non-trivial solutions of such inhomogeneous problems (subject to the usual homogeneous boundary conditions) will exist only if a certain solvability condition is satisfied and this solvability condition has been manipulated by Itoh and Nayfeh to yield the propagation equation for a second-order amplitude measure. This second-order quantity is a bilinear form in the amplitude of the solution of the small-disturbance equations of motion and the amplitude of the solution of the set of equations 'adjoint' to the small-disturbance equations.

Landahl (1982) applied Chin's method under a modified hypothesis designed to improve the accuracy of the approximation in the far field. This analysis was restricted to the case when the dispersion relation does not depend explicitly upon the propagation space coordinates. The main modification was to incorporate a convention (proposed by several other authors, including Nayfeh 1980) that the relevant cut of the complex dispersion relation in analyses of amplitude propagation is the one which renders the derivative of the frequency with respect to the wavenumber a real quantity. This cut yields complex values for both the frequency and the wavenumber and defines a mixed temporal–spatial instability problem with an unambiguous real group velocity. If proper care is taken to distinguish between each author's working definition of amplitude, one finds that the amplitude propagation equations of Itoh (1980), Nayfeh (1980) and Landahl (1982) are all compatible. We will return to the discussion of the proper definition of amplitude in §6 below.

Whitham's approach to the analysis of amplitude propagation has been a valuable complement to other methods wherever it has been applied. The main obstacle to its application to shear-flow instability waves has been the lack of a simple usable variational statement of the basic physics which application of Whitham's method requires.

Eckart (1963) and Seliger & Whitham (1968) both present general variational statements of the equation of motion of an inviscid compressible fluid. Eckart's formulation involves the use of the Lagrangian flow description. Two transformations of variables are introduced, one relating the 'present' position coordinates of a fluid particle in the undisturbed flow to the initial position coordinates of the particle, and another relating the present position coordinates in the disturbed flow to those in the undisturbed flow. No simplifications for small disturbances are introduced and the analysis, being a general one, is cumbersome. Seliger & Whitham (1968) present a variational principle in the Eulerian flow description which involves the use of 'Clebsch potentials'. That formulation, however, involves the use of variables that are difficult to relate to those usually employed in the Rayleigh stability problem.

Hayes (1970*a*) has presented the small disturbance counterpart to Eckart's equations.

In §2, we derive the equations of motion for small disturbances to an inviscid shear flow in the usual way, allowing however for the complicating effects of weak unsteadiness and non-uniformity of the disturbed flow and of non-unidirectionality of the horizontal velocity profiles in it. We take the occasion to define a reference flow flatness parameter ϵ_{rf} which will be useful in the following. In §3, we introduce a change of dependent variable which reduces the small-disturbance equations of motion to a form more easily recognized as the Euler equations of a variational principle. In §4, we present four variational principles all of which express the basic physical content of the small-disturbance equations of motion. The flow description is the Eulerian one, and the connection between the variables appearing in the variational principles and the variables normally employed in the Rayleigh instability problem is simple and direct.

In §5, we derive a conservation law for a 'bilinear wave-action density' which appears to be the closest analog of Whitham's law applicable to inviscid shear waves. We digress briefly in this section to confirm that the Euler equations of the phase-averaged variational principle lead, as one might expect, to the familiar Rayleigh stability equation. Some of the ideas of kinematic wave theory (cf. Hayes 1970*b*) are exploited in this section to obtain approximate solutions for the downstream evolution of the bilinear wave action density along a wave ray. In this section, we identify two types of near-field focusing, one of which is due to non-uniformity of the initial wavenumber distribution function for the wavetrain, while the other is due to non-uniformity of the background flow and is, in its consequences, identical to that described by Landahl (1972).

The manner in which the observed square amplitude may be defined in wavetrains capable of undergoing both dispersion and exponential amplification and decay is discussed in §6 where we propose that the bilinear wave-action density is an appropriate measure of the factor in the general formula for the observed square amplitude associated with amplitude changes by dispersion.

2. Small disturbances to a flat reference flow

Let (x_1, x_2, x_3) be a right-handed Cartesian coordinate system and let $\{\hat{\imath}_1, \hat{\imath}_2, \hat{\imath}_3\}$ be the associated set of orthonormal basis vectors. We suppose that the plane $x_2 = 0$ corresponds to a rigid impermeable plane wall. Let (U_1, U_2, U_3) be the components of the velocity field of a reference-flow solution of the inviscid incompressible uniform-density equations of motion and let P be the associated pressure field. Let $(U_1+u_1, U_2+u_2, U_3+u_3)$ and $P+p$ denote the components of the velocity field and the pressure field of a neighbouring flow solution of the equations of motion. Subtracting the reference flow equations from the neighbouring flow equations yields the equations for the disturbances, namely

$$\frac{\partial u_i}{\partial t}+U_j\frac{\partial u_i}{\partial x_j}+u_j\frac{\partial U_i}{\partial x_j}+u_j\frac{\partial u_i}{\partial x_j} = -\frac{1}{\rho}\frac{\partial p}{\partial x_i} \quad (i\in\{1,2,3\}), \tag{2.1a}$$

$$\frac{\partial u_j}{\partial x_j} = 0. \tag{2.1b}$$

We now restrict attention to the case where the reference flow is of thin shear-layer type. Specifically, if (l_1, l_2, l_3) denote typical lengthscales relative to which the reference-flow velocity field varies significantly in the x_1, x_2 and x_3 directions respectively, then we will say that the reference flow is 'flat' if the small parameter ϵ_{rf} defined by

$$\epsilon_{rf} \equiv l_2[l_1^{-2}+l_3^{-2}]^{\frac{1}{2}} \tag{2.2}$$

satisfies $$\epsilon_{rf} \ll 1.$$

Let (Q_1, Q_2, Q_3) denote velocity scales representative of the magnitudes of the components of the reference-flow velocity field. It follows from the continuity equation for the reference flow and the impermeable-wall boundary condition that

$$\begin{aligned} Q_2 &\sim l_2[Q_1/l_1+Q_3/l_3] \\ &\leqq l_2[Q_1^2+Q_3^2]^{\frac{1}{2}}[l_1^{-2}+l_3^{-2}]^{\frac{1}{2}} \\ &= \epsilon_{rf}[Q_1^2+Q_3^2]^{\frac{1}{2}}. \end{aligned}$$

For brevity, we introduce the unsubscripted symbols l and Q defined by

$$l^{-1} \equiv [l_1^{-2}+l_3^{-2}]^{\frac{1}{2}}, \tag{2.3}$$

$$Q \equiv [Q_1^2+Q_3^2]^{\frac{1}{2}}. \tag{2.4}$$

Then, by employing elementary scaling arguments, we arrive at the order-of-magnitude estimates

$$U_i \sim Q\epsilon_{rf}^{\delta_{2i}}, \tag{2.5}$$

$$\frac{\partial U_i}{\partial x_j} \sim \frac{Q}{l}\,\epsilon_{rf}^{\delta_{2i}-\delta_{2j}} \sim \frac{Q}{l_2}\,\epsilon_{rf}^{\delta_{2i}-\delta_{2j}+1}. \tag{2.6}$$

We assume that the above order-of-magnitude estimation rule may be applied inductively. That is, each additional derivative of a reference-flow velocity component introduces an additional factor ϵ_{rf}/l_2 (if the variable of differentiation is x_1 or x_3) or a factor $1/l_2$ (if the variable of differentiation is x_2) to the existing order-of-magnitude estimate of that velocity derivative. Let τ be a timescale representative of the shortest of those over which the reference-flow velocity field varies significantly. We restrict attention to the case in which

$$\tau \geqq \frac{l}{Q} \approx \frac{l_2}{\epsilon_{rf}\,Q}. \tag{2.7}$$

We then have $$\frac{\partial U_i}{\partial t} \sim \frac{Q^2}{l_2}\,\epsilon_{rf}^{1+\delta_{2i}}. \tag{2.8}$$

If in equations (2.1 *a, b*) for the disturbances, we ignore terms containing the factor $Q\epsilon_{rf}$ and define an overall disturbance-amplitude measure ϵ_q by

$$\epsilon_q^2 \equiv \frac{\max\{u_j u_j\}}{Q^2}$$

then, the disturbance equations reduce to

$$\left(\frac{\partial}{\partial t}+U_1\frac{\partial}{\partial x_1}+U_3\frac{\partial}{\partial x_3}\right)u_i+u_2\left(\frac{\partial U_1}{\partial x_2}\,\delta_{1i}+\frac{\partial U_3}{\partial x_2}\,\delta_{3i}\right) = -\frac{1}{\rho}\frac{\partial p}{\partial x_i}+O(\epsilon_{rf},\epsilon_q^2), \tag{2.9a}$$

$$\frac{\partial u_j}{\partial x_j} = 0. \tag{2.9b}$$

3. A change of dependent variables

Equations (2.9a, b) are not yet in a form that may be recognized as the Euler equations for a variational principle. Such a form may be obtained from (2.9a, b) by a simple change of variable. Let $A_i = A_i(x_i, t)$ denote a solution of the equations

$$\frac{\partial A_i}{\partial t} + U_j \frac{\partial A_i}{\partial x_j} = 0, \tag{3.1a}$$

subject to the initial conditions

$$A_i = x_i \quad \text{at } t = 0. \tag{3.1b}$$

Let $A_i + a_i$ satisfy the neighbouring flow equations

$$\frac{\partial}{\partial t}(A_i + a_i) + (U_j + u_j)\frac{\partial}{\partial x_j}(A_i + a_i) = 0, \tag{3.2a}$$

subject to the initial conditions

$$A_i + a_i = x_i \quad \text{at } t = 0. \tag{3.2b}$$

Then, by subtracting the reference-flow equations (3.1a, b) from the neighbouring flow equations (3.2a, b), the equations for the disturbances result, namely

$$\frac{\partial a_i}{\partial t} + U_j \frac{\partial a_i}{\partial x_j} + u_j \frac{\partial A_i}{\partial x_j} + u_j \frac{\partial a_i}{\partial x_j} = 0, \tag{3.3a}$$

subject to $\qquad a_i = 0 \quad \text{at } t = 0. \qquad$ (3.3b)

If we employ our reference-flow flatness hypothesis to (3.1a), then it reduces to

$$\left(\frac{\partial}{\partial t} + U_1 \frac{\partial}{\partial x_1} + U_3 \frac{\partial}{\partial x_3}\right) A_i = O(\epsilon_{\mathrm{rf}}),$$

which has the approximate solution

$$A_i = x_i - U_1 \delta_{1i} t - U_3 \delta_{3i} t + O(\epsilon_{\mathrm{rf}}), \tag{3.4}$$

[which also satisfies the initial condition (3.1b)].

Substituting (3.4) into (3.3a) and employing both the flatness and small-disturbance hypotheses wherever possible, we get

$$\left(\frac{\partial}{\partial t} + U_1 \frac{\partial}{\partial x_1} + U_3 \frac{\partial}{\partial x_3}\right) a_i + u_i - u_2 \left(\frac{\partial U_1}{\partial x_2}\delta_{1i} + \frac{\partial U_3}{\partial x_2}\delta_{3i}\right) t = O(\epsilon_{\mathrm{rf}}, \epsilon_q^2).$$

The above equation may be rearranged to yield a formula for u_i in terms of the remaining quantities, giving

$$u_i = -\left(\frac{\partial}{\partial t} + U_1 \frac{\partial}{\partial x_1} + U_3 \frac{\partial}{\partial x_3}\right) a_i - \left(\frac{\partial}{\partial t} + U_1 \frac{\partial}{\partial x_1} + U_3 \frac{\partial}{\partial x_3}\right)(a_2)\left(\frac{\partial U_1}{\partial x_2}\delta_{1i} + \frac{\partial U_3}{\partial x_2}\delta_{3i}\right) t + O(\epsilon_{\mathrm{rf}}, \epsilon_q^2)$$

or

$$u_i = -\left(\frac{\partial}{\partial t} + U_1 \frac{\partial}{\partial x_1} + U_3 \frac{\partial}{\partial x_3}\right)\left[a_i + a_2\left(\frac{\partial U_1}{\partial x_2}\delta_{1i} + \frac{\partial U_3}{\partial x_2}\delta_{3i}\right) t\right] + a_2\left(\frac{\partial U_1}{\partial x_2}\delta_{1i} + \frac{\partial U_3}{\partial x_2}\delta_{3i}\right) + O(\epsilon_{\mathrm{rf}}, \epsilon_q^2). \tag{3.5}$$

This equation may be employed to eliminate the disturbance-velocity components u_i from the small-disturbance equations of motion (2.9*a*, *b*). We find that the momentum equation (2.9*a*) reduces to the form

$$-\left(\frac{\partial}{\partial t}+U_1\frac{\partial}{\partial x_1}+U_3\frac{\partial}{\partial x_3}\right)^2\left[a_i+a_2\left(\frac{\partial U_1}{\partial x_2}\delta_{1i}+\frac{\partial U_3}{\partial x_2}\delta_{3i}\right)t\right]=-\frac{\partial}{\partial x_i}\left(\frac{p}{\rho}\right)+O(\epsilon_{\mathrm{rf}},\epsilon_q^2) \tag{3.6a}$$

and the continuity equation (2.9*b*) takes the form

$$-\left(\frac{\partial}{\partial t}+U_1\frac{\partial}{\partial x_1}+U_3\frac{\partial}{\partial x_3}\right)\left\{\frac{\partial}{\partial x_i}\left[a_i+a_2\left(\frac{\partial U_1}{\partial x_2}\delta_{1i}+\frac{\partial U_3}{\partial x_2}\delta_{3i}\right)t\right]\right\}=O(\epsilon_{\mathrm{rf}},\epsilon_q^2). \tag{3.6b}$$

The forms (3.6*a*, *b*) suggest that we regard the quantity in square brackets as a single entity. Accordingly, we introduce the shorthand

$$b_i\equiv a_i+a_2\left(\frac{\partial U_1}{\partial x_2}\delta_{1i}+\frac{\partial U_3}{\partial x_2}\delta_{3i}\right)t. \tag{3.7}$$

Another shorthand which will be useful is

$$\left(\frac{\partial}{\partial t}+U_1\frac{\partial}{\partial x_1}+U_3\frac{\partial}{\partial x_3}\right)(\ )=\mathrm{lsd}(\ ), \tag{3.8}$$

where lsd( ) is a mnemonic for linearized substantial derivative.

In view of the definition (3.7), the initial condition (3.3*b*) becomes

$$b_i=0\quad\text{at } t=0. \tag{3.9}$$

Equation (3.6*b*) now takes the form

$$\mathrm{lsd}\left(\frac{\partial b_j}{\partial x_j}\right)=O(\epsilon_{\mathrm{rf}},\epsilon_q^2),$$

which has the approximate solution

$$\frac{\partial b_j}{\partial x_j}=f(A_i)+O(\epsilon_{\mathrm{rf}},\epsilon_q^2),$$

where f is a differentiable function of three arguments. Applying the initial condition (3.9), we conclude that f must be zero. In our shorthand notation, the momentum and continuity equations now become

$$\mathrm{lsd}^2(b_i)-\frac{\partial}{\partial x_i}\left(\frac{p}{\rho}\right)=O(\epsilon_{\mathrm{rf}},\epsilon_q^2),\quad \frac{\partial b_j}{\partial x_j}=O(\epsilon_{\mathrm{rf}},\epsilon_q^2). \tag{3.10a, b}$$

In summary, equations (3.10*a*, *b*) are equivalent to the more familiar equations of motion (2.9*a*, *b*) for small disturbances to a flat reference flow. The former follows from the latter by the change of variable

$$u_i=-\mathrm{lsd}(b_i)+b_2\left(\frac{\partial U_1}{\partial x_2}\delta_{1i}+\frac{\partial U_3}{\partial x_2}\delta_{3i}\right), \tag{3.11}$$

[which is a rewriting of (3.5) in the shorthand defined by (3.7) and (3.8)].

4. Variational principles

4.1. *Variational principles in physical variables*

Let D denote the region of three-dimensional space interior to the right circular cylinder $\mathscr{S}$ whose boundaries are the planes $x_2 = 0$ and $x_2 = h$ and the cylindrical surface $x_1^2+x_3^2 = R^2$, where h and R are given lengths. Let t_0 and t_1 be given times. Consider the variational principal

$$\delta \iiint_D \int_{t_0}^{t_1} \left[\tfrac{1}{2}\,\mathrm{lsd}\,(b_j)\,\mathrm{lsd}\,(b_j)-\frac{p}{\rho}\frac{\partial b_j}{\partial x_j}\right] \mathrm{d}t\,\mathrm{d}x_1\,\mathrm{d}x_2\,\mathrm{d}x_3 = O(\epsilon_{\mathrm{rf}}, \epsilon_q^3), \tag{4.1}$$

in which the three components (b_1, b_2, b_3) and p are regarded as independently variable. We suppose that b_j is given prescribed values on the surface $x_1^2+x_3^2 = R^2$ and that b_2 is given prescribed values on the surfaces $x_2 = 0$ and $x_2 = h$. We also suppose that b_j is given prescribed values at the temporal endpoints t_0 and t_1. The variations of b_j must then be zero on those surfaces and at those times where its values are prescribed. Carrying out the usual operations of variational calculus including integration by parts and use of arbitrariness of the test functions δb_j and δp throughout the interior of D and using the flatness hypothesis wherever possible to ignore the derivatives of U_1 and U_3 with respect to x_1 and x_3, and t, we find that the Euler equation corresponding to independent variations with respect to b_j is the momentum equation (3.10*a*) and the Euler equation corresponding to independent variations with respect to p is the continuity equation (3.10*b*).

The variational principle (4.1) is the most direct variational statement of (3.10*a*, *b*), but it is not the only one. Another principle which expresses the same physics but which we will find more useful when we substitute trial solutions in the form of slowly varying wavetrains, which may experience exponential growth or decay along the ray, is the bilinear variational principle

$$\delta \iiint_D \int_{t_0}^{t_1} \left[\mathrm{lsd}\,(b_j)\,\mathrm{lsd}\,(\mathscr{b}_j)-\left(\frac{p}{\rho}\right)\frac{\partial \mathscr{b}_j}{\partial x_j}-\left(\frac{\mathscr{p}}{\rho}\right)\frac{\partial b_j}{\partial x_j}\right] \mathrm{d}t\,\mathrm{d}x_1\,\mathrm{d}x_2\,\mathrm{d}x_3 = O(\epsilon_{\mathrm{rf}}, \epsilon_q^3), \tag{4.2}$$

in which the eight quantities (b_1, b_2, b_3), $(\mathscr{b}_1, \mathscr{b}_2, \mathscr{b}_3)$, p, and $\mathscr{p}$ are all treated as independently variable and where the surfaces and times where b_j and $\mathscr{b}_j$ take prescribed values are the same as those listed earlier for b_j in the variational principle (4.1).

The Euler equations of the principle (4.2) corresponding to independent variations of the script-type variables $\delta(\mathscr{b}_j)$ and $\delta(\mathscr{p})$ are equations (3.10*a*) and (3.10*b*) respectively. The Euler equations of (4.2) corresponding to independent variations of the italic variables $\delta(b_j)$ and $\delta(p)$ are

$$\mathrm{lsd}^2\,\mathscr{b}_i-\frac{\partial}{\partial x_i}\left(\frac{\mathscr{p}}{\rho}\right) = O(\epsilon_{\mathrm{rf}}, \epsilon_q^2), \quad \frac{\partial}{\partial x_j}\left(\frac{\mathscr{b}_j}{\rho}\right) = O(\epsilon_{\mathrm{rf}}, \epsilon_q^2) \tag{4.3a, b}$$

respectively, which are equivalent to (3.10*a*, *b*).

There are many circumstances in which a single equation for a single unknown is more useful than a set of equations for a set of unknowns. Thus, if one starts with the conventional form (2.9*a*, *b*) of the small-disturbance equations of motion, it is possible to operate on the original set of four equations to eliminate the variables u_1, u_3, and p, leaving a single higher-order equation for u_2. In a like manner, we may operate on the system (3.10*a*, *b*) to eliminate b_1, b_3, and p leaving a single higher-order equation for b_2. The latter equation for b_2 may be expressed in variational form in at least two ways. These two variational principles for the b_2 equation are equivalent in the same sense that the variational principles (4.1) and (4.2) are (i.e. one involves the use of both 'script' and 'italic' test functions while the other involves only 'italic' functions).

Thus if the operator

$$\delta_{i1}\frac{\partial}{\partial x_1}+\delta_{i3}\frac{\partial}{\partial x_3}$$

is applied to (3.10*a*) and if (3.10*b*) is substituted in the form

$$\frac{\partial b_1}{\partial x_1}+\frac{\partial b_3}{\partial x_3}=-\frac{\partial b_2}{\partial x_2}+O(\epsilon_{\mathrm{rf}},\epsilon_q^2),$$

we get

$$\mathrm{lsd}^2\left(-\frac{\partial b_2}{\partial x_2}\right)-\left(\frac{\partial^2}{\partial x_1^2}+\frac{\partial^2}{\partial x_3^2}\right)\frac{p}{\rho}=O(\epsilon_{\mathrm{rf}},\epsilon_q^2). \tag{4.4}$$

Eliminating the pressure between this equation and the middle component of (3.10*a*), i.e.

$$\mathrm{lsd}^2\,b_2-\frac{\partial}{\partial x_2}\left(\frac{p}{\rho}\right)=O(\epsilon_{\mathrm{rf}},\epsilon_q^2), \tag{4.5}$$

we get

$$\frac{\partial}{\partial x_j}\left[\mathrm{lsd}^2\left(\frac{\partial b_2}{\partial x_j}\right)\right]=O(\epsilon_{\mathrm{rf}},\epsilon_q^2). \tag{4.6}$$

Consider the variational principle

$$\delta\iiint_D\int_{t_0}^{t_1}\mathrm{lsd}\left(\frac{\partial b_2}{\partial x_j}\right)\mathrm{lsd}\left(\frac{\partial b_2}{\partial x_j}\right)\mathrm{d}t\,\mathrm{d}x_1\,\mathrm{d}x_2\,\mathrm{d}x_3=O(\epsilon_{\mathrm{rf}},\epsilon_q^3), \tag{4.7}$$

in which $\partial b_2/\partial x_j$ is given prescribed values on the cylindrical surface $x_1^2+x_3^2=R^2$ and the temporal endpoints t_0 and t_1, and b_2 is given prescribed values on the cylindrical surface $x_1^2+x_3^2=R^2$ and on the planes $x_2=0$ and $x_2=h$. Under these conditions, the Euler equation corresponding to arbitrary variations of b_2 in the interior of D is equation (4.6).

The bilinear counterpart of (4.7) is the principle

$$\delta\iiint_D\int_{t_0}^{t_1}\mathrm{lsd}\left(\frac{\partial b_2}{\partial x_j}\right)\mathrm{lsd}\left(\frac{\partial \mathscr{b}_2}{\partial x_j}\right)\mathrm{d}t\,\mathrm{d}x_1\,\mathrm{d}x_2\,\mathrm{d}x_3=O(\epsilon_{\mathrm{rf}},\epsilon_q^3), \tag{4.8}$$

in which $\partial b_2/\partial x_j$ and $\partial \mathscr{b}_2/\partial x_j$ are given prescribed values on the cylinder $x_1^2+x_3^2=R^2$ and at the temporal endpoints t_0 and t_1, while b_2 and $\mathscr{b}_2$ are given prescribed values on the cylinder $x_1^2+x_3^2=R^2$ and on the planes $x_2=0$ and $x_2=h$. The Euler equation corresponding to independent variations of $\mathscr{b}_2$ is (4.6) while the Euler equation corresponding to independent variations of b_2 is

$$\frac{\partial}{\partial x_j}\left[\mathrm{lsd}^2\left(\frac{\partial \mathscr{b}_2}{\partial x_j}\right)\right]=O(\epsilon_{\mathrm{rf}},\epsilon_q^2). \tag{4.9}$$

4.2. *Slowly varying wavetrains*

Let θ be a scalar function (possibly complex valued) of the real variables x_1, x_3 and t. We will say that H is a slowly varying quantity relative to a wavetrain with phase function θ if the second term on the right-hand side of each of the identities

$$\frac{\partial}{\partial t}(H\mathrm{e}^{\pm \mathrm{i}\theta})=\pm\mathrm{i}\frac{\partial\theta}{\partial t}H\mathrm{e}^{\pm\mathrm{i}\theta}+\frac{\partial H}{\partial t}\mathrm{e}^{\pm\mathrm{i}\theta}, \tag{4.10}$$

$$\frac{\partial}{\partial x_i}(H\mathrm{e}^{\pm \mathrm{i}\theta})=\pm\mathrm{i}\frac{\partial\theta}{\partial x_i}H\mathrm{e}^{\pm\mathrm{i}\theta}+\frac{\partial H}{\partial x_i}\mathrm{e}^{\pm\mathrm{i}\theta},\quad i\in\{1,3\} \tag{4.11}$$

is of uniformly small magnitude compared to the magnitude of the first term. We will call the product $H\exp(\pm i\theta)$ a slowly varying wavetrain only if two conditions are met:

(i) H is slowly varying in the above sense;

(ii) All of the various higher-order partial derivatives of H and θ with respect to the variables x_1, x_3 and t are slowly varying in the above sense.

We now substitute test functions in the form of slowly varying wavetrains into the bilinear variational principle (4.2). Specifically, we let

$$b_i = 2\,\mathrm{Re}\{\tilde{b}_i e^{i\theta}\} = \tilde{b}_i e^{i\theta} + \tilde{b}_i^* e^{-i\theta^*}, \tag{4.12a}$$

$$p = 2\,\mathrm{Re}\{\tilde{p} e^{i\theta}\} = \tilde{p} e^{i\theta} + \tilde{p}^* e^{-i\theta^*}, \tag{4.12b}$$

$$\mathscr{b}_i = 2\,\mathrm{Re}\{\tilde{\mathscr{b}}_i e^{-i\theta}\} = \tilde{\mathscr{b}}_i e^{-i\theta} + \tilde{\mathscr{b}}_i^* e^{i\theta^*}, \tag{4.12c}$$

$$\mathscr{p} = 2\,\mathrm{Re}\{\tilde{\mathscr{p}} e^{-i\theta}\} = \tilde{\mathscr{p}} e^{-i\theta} + \tilde{\mathscr{p}}^* e^{i\theta^*}, \tag{4.12d}$$

in which all of the quantities with an overhead tilde are (in general) complex-valued functions of the real variables x_1, x_2, x_3 and t and the asterisk denotes complex-conjugation.

Let ϵ_w be the wavetrain non-uniformity parameter which measures the size of the second term on the right-hand side of (4.10) or (4.11) compared to that of the first. Then, the dominant parts of the various factors in square brackets in (4.2) are

$$\mathrm{lsd}\,(b_j) = \tilde{b}_j\,\mathrm{lsd}\,(i\theta)\,e^{i\theta} + \tilde{b}_j^*\,\mathrm{lsd}\,(-i\theta^*)\,e^{-i\theta^*} + O(\epsilon_w),$$

$$\mathrm{lsd}\,(\mathscr{b}_j) = \tilde{\mathscr{b}}_j\,\mathrm{lsd}\,(-i\theta)\,e^{-i\theta} + \tilde{\mathscr{b}}_j^*\,\mathrm{lsd}\,(i\theta^*)\,e^{i\theta^*} + O(\epsilon_w),$$

$$\frac{\partial b_j}{\partial x_j} = \left[\tilde{b}_1 \frac{\partial(i\theta)}{\partial x_1} + \frac{\partial \tilde{b}_2}{\partial x_2} + \tilde{b}_3 \frac{\partial(i\theta)}{\partial x_3}\right] e^{i\theta} + \left[\tilde{b}_1^* \frac{\partial(-i\theta^*)}{\partial x_1} + \frac{\partial \tilde{b}_2^*}{\partial x_2} + \tilde{b}_3^* \frac{\partial(-i\theta^*)}{\partial x_3}\right] e^{-i\theta^*} + O(\epsilon_w),$$

$$\frac{\partial \mathscr{b}_j}{\partial x_j} = \left[\tilde{\mathscr{b}}_1 \frac{\partial(-i\theta)}{x_1} + \frac{\partial \tilde{\mathscr{b}}_2}{\partial x_2} + \tilde{\mathscr{b}}_3 \frac{\partial(-i\theta)}{\partial x_3}\right] e^{-i\theta} + \left[\tilde{\mathscr{b}}_1^* \frac{\partial(i\theta^*)}{\partial x_1} + \frac{\partial \tilde{\mathscr{b}}_2^*}{\partial x_2} + \tilde{\mathscr{b}}_3^* \frac{\partial(i\theta^*)}{\partial x_3}\right] e^{i\theta^*} + O(\epsilon_w).$$

It follows that

$$\mathrm{lsd}\,(b_j)\,\mathrm{lsd}\,(\mathscr{b}_j) = 2\,\mathrm{Re}\{\tilde{b}_j \tilde{\mathscr{b}}_j (\mathrm{lsd}\,\theta)^2\} - |\mathrm{lsd}\,\theta|^2\, 2\,\mathrm{Re}\{\tilde{b}_j \tilde{\mathscr{b}}_j^*\, e^{i(\theta+\theta^*)}\} + O(\epsilon_w), \tag{4.13}$$

$$\frac{p}{\rho}\frac{\partial \mathscr{b}_j}{\partial x_j} = 2\,\mathrm{Re}\left\{\frac{\tilde{p}}{\rho}\left[\tilde{\mathscr{b}}_1 \frac{\partial(-i\theta)}{\partial x_1} + \frac{\partial \tilde{\mathscr{b}}_2}{\partial x_2} + \tilde{\mathscr{b}}_3 \frac{\partial(-i\theta)}{\partial x_3}\right]\right\}$$
$$+ 2\,\mathrm{Re}\left\{\frac{p}{\rho}\left[\tilde{\mathscr{b}}_1^* \frac{\partial(i\theta^*)}{\partial x_1} + \frac{\partial \tilde{\mathscr{b}}_2^*}{\partial x_2} + \tilde{\mathscr{b}}_3^* \frac{\partial(i\theta^*)}{\partial x_3}\right] e^{i(\theta+\theta^*)}\right\} + O(\epsilon_w), \tag{4.14}$$

$$\frac{\mathscr{p}}{\rho}\frac{\partial b_j}{\partial x_j} = 2\,\mathrm{Re}\left\{\frac{\tilde{\mathscr{p}}}{\rho}\left[\tilde{b}_1 \frac{\partial(i\theta)}{\partial x_1} + \frac{\partial \tilde{b}_2}{\partial x_2} + \tilde{b}_3 \frac{\partial(i\theta)}{\partial x_3}\right]\right\}$$
$$+ 2\,\mathrm{Re}\left\{\frac{\tilde{\mathscr{p}}}{\rho}\left[\tilde{b}_1^* \frac{\partial(-i\theta^*)}{\partial x_1} + \frac{\partial \tilde{b}_2^*}{\partial x_2} + \tilde{b}_3^* \frac{\partial(-i\theta^*)}{\partial x_3}\right] e^{-i(\theta+\theta^*)}\right\} + O(\epsilon_w). \tag{4.15}$$

We will refer to the quantity in square brackets in (4.2) as the 'bilinear Lagrangian density' and denote it by the symbol L. If we define a phase-averaging operator $\langle\ \rangle$ by

$$\langle\ \rangle \equiv \frac{1}{2\pi}\int_{\theta_r=0}^{2\pi} (\ )\,d\theta_r,$$

then, we may calculate $\langle L\rangle$ from (4.13), (4.14) and (4.15) and obtain

$$\langle L\rangle = 2\,\mathrm{Re}\{\tilde{L}\},$$

where

$$\tilde{L} \equiv \tilde{b}_j\tilde{\ell}_j(\mathrm{lsd}\,\theta)^2 - \frac{\tilde{p}}{\rho}\left[\tilde{\ell}_1\frac{\partial(-\mathrm{i}\theta)}{\partial x_1}+\frac{\partial\tilde{\ell}_2}{\partial x_2}+\tilde{\ell}_3\frac{\partial(-\mathrm{i}\theta)}{\partial x_3}\right] - \frac{\tilde{\not{p}}}{\rho}\left[\tilde{b}_1\frac{\partial(\mathrm{i}\theta)}{\partial x_1}+\frac{\partial\tilde{b}_2}{\partial x_2}+\tilde{b}_3\frac{\partial(\mathrm{i}\theta)}{\partial x_3}\right]+O(\epsilon_{\mathrm{w}}). \quad (4.16)$$

4.3. *Phase-averaged variational principal*

Following the general scenario described by Whitham (1974, p. 393), we now propose the 'phase-averaged bilinear variational principle'

$$\delta\iiint_D\int_{t_0}^{t_1} 2\,\mathrm{Re}\{\tilde{L}\}\,\mathrm{d}t\,\mathrm{d}x_1\,\mathrm{d}x_2\,\mathrm{d}x_3 = O(\epsilon_{\mathrm{rf}},\epsilon_q^3,\epsilon_{\mathrm{w}}), \quad (4.17)$$

where

$$\tilde{L} = \tilde{L}\left(\tilde{b}_i,\frac{\partial\tilde{b}_2}{\partial x_2},\tilde{\ell}_i,\frac{\partial\tilde{\ell}_2}{\partial x_2},\tilde{p},\tilde{\not{p}},-\frac{\partial\theta}{\partial t},\frac{\partial\theta}{\partial x_1},\frac{\partial\theta}{\partial x_3},x_i,t\right), \quad (4.18)$$

is defined by (4.16). The explicit dependence of $\tilde{L}$ on x_i and t is due to the dependence of U_1 and U_3 on those variables and the fact that U_1 and U_3 appear in the definition (3.8) of the operator $\underset{\sim}{\mathrm{lsd}}(\ )$.

In (4.17), we treat $\tilde{b}_i$, $\tilde{\ell}_i$, $\tilde{p}$, $\tilde{\not{p}}$, and θ as independently variable. We suppose that $\tilde{b}_2$ and $\tilde{\ell}_2$ are given prescribed values on the planes $x_2 = 0$ and $x_2 = h$ and that θ is given prescribed values on the cylinder $x_1^2+x_3^2 = R^2$ (for all t satisfying $t_0 \leqslant t \leqslant t_1$) and at the temporal endpoints t_0 and t_1 (for all $x_i \in D$).

Before writing the Euler equations corresponding to the above variational principle, we will establish a notation convention to avoid ambiguity in the interpretation of partial-derivative symbols. The convention will apply whenever a derivative of a composite function (i.e. a function of a function) is taken and both an 'implicit' interpretation (involving a chain-rule expansion) and an 'explicit' interpretation (involving only differentiation with respect to the explicit dependence of the operand) is possible. We will avoid differentiation of 'functions of functions of functions' or other examples of multiple composition.

Thus, subscripts will be employed to denote 'explicit' partial derivatives, e.g.

$$\tilde{L}_{x_1},\quad \tilde{L}_{x_3},\quad \tilde{L}_t,$$

and 'complete' partial-derivative symbols will be employed to denote 'implicit' partial derivatives involving the full-chain-rule expansion, e.g.

$$\frac{\partial\tilde{L}}{\partial x_1},\quad \frac{\partial\tilde{L}}{\partial x_3},\quad \frac{\partial\tilde{L}}{\partial t}.$$

The same convention will be retained in writing derivatives of the 'dispersion relation function' (cf. §4 below).

The Euler equations for the variational principle (4.17) may be written immediately.

Some of the formulas which arise during their derivation will be useful for later reference, however, so we include them here. We have

$$\delta\tilde{L} = \tilde{L}_{\tilde{b}_j}\,\delta\tilde{b}_j + \tilde{L}_{(\partial\tilde{b}_2/\partial x_2)}\,\delta\left(\frac{\partial\tilde{b}_2}{\partial x_2}\right) + \tilde{L}_{\tilde{\mathscr{E}}_j}\,\delta\tilde{\mathscr{E}}_j$$
$$+ \tilde{L}_{(\partial\tilde{\mathscr{E}}_2/\partial x_2)}\,\delta\left(\frac{\partial\tilde{\mathscr{E}}_2}{\partial x_2}\right) + \tilde{L}_{\tilde{p}}\,\delta\tilde{p} + \tilde{L}_{\tilde{\not{p}}}\,\delta\tilde{\not{p}}$$
$$+ \tilde{L}_{(-\partial\theta/\partial t)}\,\delta\left(\frac{-\partial\theta}{\partial t}\right) + \tilde{L}_{(\partial\theta/\partial x_1)}\,\delta\left(\frac{\partial\theta}{\partial x_1}\right) + \tilde{L}_{(\partial\theta/\partial x_3)}\,\delta\left(\frac{\partial\theta}{\partial x_3}\right)$$

or

$$\delta\tilde{L} = \frac{\partial}{\partial x_2}\,[\tilde{L}_{(\partial\tilde{b}_2/\partial x_2)}\,\delta\tilde{b}_2 + \tilde{L}_{(\partial\tilde{\mathscr{E}}_2/\partial x_2)}\,\delta\tilde{\mathscr{E}}_2] + \left[\tilde{L}_{\tilde{b}_j} - \delta_{2j}\,\frac{\partial}{\partial x_2}\,(\tilde{L}_{(\partial\tilde{b}_2/\partial x_2)})\right]\delta\tilde{b}_j$$
$$+ \tilde{L}_{\tilde{p}}\,\delta\tilde{p} + \left[\tilde{L}_{\tilde{\mathscr{E}}_j} - \delta_{2j}\,\frac{\partial}{\partial x_2}\,(\tilde{L}_{(\partial\tilde{\mathscr{E}}_2/x_2)})\right]\delta\tilde{\mathscr{E}}_j + \tilde{L}_{\tilde{\not{p}}}\,\delta\tilde{\not{p}}$$
$$+ \tilde{L}_{(-\partial\theta/\partial t)}\,\delta\left(-\frac{\partial\theta}{\partial t}\right) + \tilde{L}_{(\partial\theta/\partial x_1)}\,\delta\left(\frac{\partial\theta}{\partial x_1}\right) + \tilde{L}_{(\partial\theta/\partial x_3)}\,\delta\left(\frac{\partial\theta}{\partial x_3}\right). \qquad (4.19)$$

The quantity in the last line of (4.19) may be rewritten according to the identity

$$\tilde{L}_{(-\partial\theta/\partial t)}\,\delta\left(-\frac{\partial\theta}{\partial t}\right) + \tilde{L}_{(\partial\theta/\partial x_1)}\,\delta\left(\frac{\partial\theta}{\partial x_1}\right) + \tilde{L}_{(\partial\theta/\partial x_3)}\,\delta\left(\frac{\partial\theta}{\partial x_3}\right)$$
$$= \frac{\partial}{\partial t}\left[\frac{1}{\epsilon_w}\,\tilde{L}_{(\partial\theta/\partial t)}\,\delta(-\epsilon_w\,\theta)\right] + \frac{\partial}{\partial x_1}\left[\frac{1}{\epsilon_w}\,\tilde{L}_{(\partial\theta/\partial x_1)}\,\delta(\epsilon_w\,\theta)\right]$$
$$+ \frac{\partial}{\partial x_3}\left[\frac{1}{\epsilon_w}\,\tilde{L}_{(\partial\theta/\partial x_3)}\,\delta(\epsilon_w\,\theta)\right] - \left[\frac{-1}{\epsilon_w}\,\frac{\partial}{\partial t}\,(\tilde{L}_{(-\partial\theta/\partial t)})\right.$$
$$\left. + \frac{1}{\epsilon_w}\,\frac{\partial}{\partial x_1}\,(\tilde{L}_{(\partial\theta/\partial x_1)}) + \frac{1}{\epsilon_w}\,\frac{\partial}{\partial x_3}\,(\tilde{L}_{(\partial\theta/\partial x_3)})\right]\delta(\epsilon_w\,\theta). \qquad (4.20)$$

The rescaling of the factors on the right-hand side of (4.20) has been made so that the individual terms in the coefficient of $\delta(\epsilon_w\,\theta)$ will be nominally bounded as $\epsilon_w \to 0$. Without the factor ϵ_w^{-1}, terms such as

$$\frac{\partial}{\partial t}\,(\tilde{L}_{(\partial\theta/\partial t)})$$

(which consist of a propagation space derivative of a slowly varying quantity) would vanish as $\epsilon_w \to 0$. In such a situation, the Euler equation corresponding to independent variations of θ would reduce to a triviality, since it would then state that the sum of three terms which are individually zero is zero.

The Euler equations for the variational principle (4.17) now become

$$\delta\tilde{\mathscr{E}}_j\colon\quad \tilde{L}_{\tilde{\mathscr{E}}_j} - \delta_{2j}\,\frac{\partial}{\partial x_2}\,(\tilde{L}_{(\partial\tilde{\mathscr{E}}_2/\partial x_2)}) = O(\epsilon_{\mathrm{rf}}, \epsilon_q^2, \epsilon_w), \qquad (4.21)$$

$$\delta\tilde{\not{p}}\colon\quad \tilde{L}_{\tilde{\not{p}}} = O(\epsilon_{\mathrm{rf}}, \epsilon_q^2, \epsilon_w), \qquad (4.22)$$

$$\delta\tilde{b}_j: \quad \tilde{L}_{\tilde{b}_j} - \delta_{2j}\frac{\partial}{\partial x_2}(\tilde{L}_{(\partial\tilde{b}_2/\partial x_2)}) = O(\epsilon_{\mathrm{rf}}, \epsilon_q^2, \epsilon_{\mathrm{w}}), \tag{4.23}$$

$$\delta\tilde{p}: \quad \tilde{L}_{\tilde{p}} = O(\epsilon_{\mathrm{rf}}, \epsilon_q^2, \epsilon_{\mathrm{w}}), \tag{4.24}$$

$$\delta(\epsilon_{\mathrm{w}}\theta): \quad \frac{1}{\epsilon_{\mathrm{w}}}\left[-\frac{\partial}{\partial t}(\tilde{L}_{(-\partial\theta/\partial t)}) + \frac{\partial}{\partial x_1}(\tilde{L}_{(\partial\theta/\partial x_1)}) + \frac{\partial}{\partial x_3}(\tilde{L}_{(\partial\theta/\partial x_3)})\right] = O(\epsilon_{\mathrm{rf}}, \epsilon_q^3, \epsilon_{\mathrm{w}}). \tag{4.25}$$

4.4. *The homogeneous boundary-value problem*

From the above Euler equations and the bilinear dependence of $\tilde{L}$ upon the script and the italic systems of variables, we note that the system (4.21), (4.22) [after the specific formula (4.16) for $\tilde{L}$ has been inserted] will not involve any of the script variables. The system (4.21), (4.22) is, therefore, uncoupled from the system (4.23), (4.24). Furthermore, the only derivatives appearing in (4.21), (4.22) are derivatives with respect to x_2. Of the four variables involved, namely

$$\tilde{b}_1, \quad \tilde{b}_2, \quad \tilde{b}_3, \quad \tilde{p},$$

the variables $\tilde{b}_1$ and $\tilde{b}_3$ appear only in undifferentiated form and may be eliminated algebraically. When such an elimination is carried out, a two-by-two matrix differential system for the variables $\tilde{b}_2$ and $\tilde{p}$ results. Either of these two variables may be eliminated by cross-differentiation leaving a second-order ordinary differential equation for the other. The equation for $\tilde{b}_2$, for example, is found to be

$$\frac{\partial}{\partial x_2}\left[(\mathrm{lsd}\,\theta)^2\frac{\partial\tilde{b}_2}{\partial x_2}\right] - \left[\left(\frac{\partial\theta}{\partial x_1}\right)^2 + \left(\frac{\partial\theta}{\partial x_3}\right)^2\right](\mathrm{lsd}\,\theta)^2\,\tilde{b}_2 = O(\epsilon_{\mathrm{rf}}, \epsilon_q^2, \epsilon_{\mathrm{w}}). \tag{4.26}$$

If, instead of starting with the variational principle (4.2), we had started from the alternative variational principle (4.8), then, after insertion of a trial solution in the form of a slowly varying wavetrain and phase averaging, the resulting Euler equation would have been the above equation.

Indeed, if the slowly varying wavetrain (4.12*a*) (with $i = 2$) were inserted into the small disturbance equation of motion in the form (4.6), the same equation would follow. The above equation is one way of writing the Rayleigh stability equation.

Introducing the notation for the (complex) wavenumber vector

$$\boldsymbol{k} = k_1\hat{\imath}_1 + k_3\hat{\imath}_3 \equiv \hat{\imath}_1\frac{\partial\theta}{\partial x_1} + \hat{\imath}_3\frac{\partial\theta}{\partial x_3}, \tag{4.27}$$

and the (complex) frequency

$$\omega \equiv -\frac{\partial\theta}{\partial t}, \tag{4.28}$$

the above equation for $\tilde{b}_2$ becomes

$$\frac{\partial}{\partial x_2}\left[(-\omega + \boldsymbol{k}\cdot\boldsymbol{U})^2\frac{\partial\tilde{b}_2}{\partial x_2}\right] - \boldsymbol{k}\cdot\boldsymbol{k}(-\omega + \boldsymbol{k}\cdot\boldsymbol{U})^2\,\tilde{b}_2 = O(\epsilon_{\mathrm{rf}}, \epsilon_q^2, \epsilon_{\mathrm{w}}),$$

which is similar to the 'displacement' form of the Rayleigh stability equation employed, for example, by L. N. Howard in the derivation of the 'semicircle theorem' of parallel-flow stability theory (cf. Drazin & Howard 1966, equation (2.23)). The only difference between the above equation and the one employed by Howard is that we have here allowed for non-unidirectionality of the horizontal velocity profile.

A completely analogous sequence of results may be derived from the set of Euler equations (4.23), (4.24). Specifically, we find that the system (4.23), (4.24) involves only the script variables. A two-by-two matrix differential system for $\tilde{\ell}_2$ and $\tilde{\mathscr{p}}$ may be obtained by algebraic elimination of $\tilde{\ell}_1$ and $\tilde{\ell}_3$. Finally, a single equation for $\tilde{\ell}_2$ follows by cross-differentiation and is found to be identical with (4.26) with $\tilde{b}_2$ replaced by $\tilde{\ell}_2$.

We now consider solutions for $\tilde{b}_2$ and $\tilde{\ell}_2$ subject to the homogeneous boundary conditions

$$(\tilde{b}_2, \tilde{\ell}_2) = (0,0) \quad \text{at } x_2 = 0,\, x_2 = h. \tag{4.29}$$

Since $\tilde{b}_2$ (or $\tilde{\ell}_2$) satisfies a homogeneous differential equation subject to homogeneous boundary conditions, non-trivial solutions will exist only for special combinations of the parameters. We have, therefore, a single dispersion relation applicable to the boundary-value problem for either $\tilde{b}_2$ or $\tilde{\ell}_2$ of the form

$$F(\omega, k_1, k_3, x_1, x_3, t) = O(\epsilon_{\mathrm{rf}}, \epsilon_q, \epsilon_{\mathrm{w}}). \tag{4.30}$$

In principle, (4.30) may be solved implicitly for ω to give

$$\omega = \Omega(k_1, k_3, x_1, x_3, t) + O(\epsilon_{\mathrm{rf}}, \epsilon_q, \epsilon_{\mathrm{w}}), \tag{4.31}$$

which will be more useful in what follows.

5. Conservation of wave-action density

We let the symbol $\mathscr{L}$ denote the cross-space integral of $\tilde{L}$, i.e.

$$\mathscr{L} \equiv \int_0^h \tilde{L}\,\mathrm{d}x_2. \tag{5.1}$$

We can show that the order of magnitude of $\mathscr{L}$, evaluated for solutions of the Euler equations (4.21)–(4.24) subject to the homogeneous boundary conditions (4.29), is as small or smaller than terms already neglected in the following way. From the bilinear dependence of $\tilde{L}$ upon the script and italic variables [cf. (4.16)], we have

$$\tilde{L} = \frac{1}{2}\left\{\tilde{L}_{\tilde{b}_j}\tilde{b}_j + \tilde{L}_{\tilde{\ell}_j}\tilde{\ell}_j + \tilde{L}_{(\partial\tilde{b}_2/\partial x_2)}\frac{\partial\tilde{b}_2}{\partial x_2} + \tilde{L}_{(\partial\tilde{\ell}_2/\partial x_2)}\frac{\partial\tilde{\ell}_2}{\partial x_2} + \tilde{L}_{\tilde{p}}\tilde{p} + \tilde{L}_{\tilde{\mathscr{p}}}\tilde{\mathscr{p}}\right\} + O(\epsilon_{\mathrm{w}})$$

or

$$\tilde{L} = \frac{1}{2}\left\{\frac{\partial}{\partial x_2}[\tilde{L}_{(\partial\tilde{b}_2/\partial x_2)}\tilde{b}_2 + \tilde{L}_{(\partial\tilde{\ell}_2/\partial x_2)}\tilde{\ell}_2] + \left[\tilde{L}_{\tilde{b}_j} - \delta_{2j}\frac{\partial}{\partial x_2}(\tilde{L}_{(\partial\tilde{b}_2/\partial x_2)})\right]\tilde{b}_j\right.$$
$$\left. + \tilde{L}_{\tilde{p}}\tilde{p} + \left[\tilde{L}_{\tilde{\ell}_j} - \delta_{2j}\frac{\partial}{\partial x_2}(\tilde{L}_{(\partial\tilde{\ell}_2/\partial x_2)})\right]\tilde{\ell}_j + \tilde{L}_{\tilde{\mathscr{p}}}\tilde{\mathscr{p}}\right\} + O(\epsilon_{\mathrm{w}}).$$

It follows from the Euler equations (4.21)–(4.24) that the quantities in the second line and the coefficient of $\tilde{b}_j$ in the first are $O(\epsilon_{\mathrm{rf}}, \epsilon_q^3, \epsilon_{\mathrm{w}})$. Integrating over the 'cross'-space and applying the homogeneous boundary conditions (4.29), we obtain

$$\int_0^h \tilde{L}\,\mathrm{d}x_2 \equiv \mathscr{L} = O(\epsilon_{\mathrm{rf}}, \epsilon_q^3, \epsilon_{\mathrm{w}}) \tag{5.2}$$

as stated. Let the symbol $\delta^{\dagger}(\;)$ denote variations within the family of solutions to the full two-point boundary-value problem. Within this family, the Euler equations (4.21)–(4.24) are satisfied and the coefficients of $\delta\tilde{b}_j$, $\delta\tilde{\ell}_j$, $\delta\tilde{p}$, and $\delta\tilde{\mathscr{p}}$ in the identity

(4.19) are of higher order under approximations already stated. Integrating this (simplified) identity over the cross-space and applying the boundary conditions (4.29), we get

$$\delta^\dagger \mathscr{L} = \mathscr{L}_\omega \delta^\dagger \omega + \mathscr{L}_{k_1} \delta^\dagger k_1 + \mathscr{L}_{k_3} \delta^\dagger k_3 + O(\epsilon_{\mathrm{rf}}, \epsilon_q^3, \epsilon_{\mathrm{w}}),$$

which, in view of (5.2), implies that

$$\mathscr{L}_\omega \delta^\dagger \omega + \mathscr{L}_{k_1} \delta^\dagger k_1 + \mathscr{L}_{k_3} \delta^\dagger k_3 = O(\epsilon_{\mathrm{rf}}, \epsilon_q^3, \epsilon_{\mathrm{w}}).$$

Applying the $\delta^\dagger(\ )$ operator to (4.31), we get

$$\delta^\dagger \omega = \Omega_{k_1} \delta^\dagger k_1 + \Omega_{k_3} \delta^\dagger k_3 + O(\epsilon_{\mathrm{rf}}, \epsilon_q, \epsilon_{\mathrm{w}}).$$

Substituting into the preceding equation to eliminate $\delta^\dagger(\omega)$ results in

$$(\mathscr{L}_{k_1} + \Omega_{k_1} \mathscr{L}_\omega) \delta^\dagger k_1 + (\mathscr{L}_{k_3} + \Omega_{k_3} \mathscr{L}_\omega) \delta^\dagger k_3 = O(\epsilon_{\mathrm{rf}}, \epsilon_q^3, \epsilon_{\mathrm{w}}).$$

From arbitrariness of $\delta^\dagger k_1$ and $\delta^\dagger k_3$ we have

$$\mathscr{L}_{k_j} = -\Omega_{k_j} \mathscr{L}_\omega + O(\epsilon_{\mathrm{rf}}, \epsilon_q^3, \epsilon_{\mathrm{w}}), \quad j \in \{1, 3\} \tag{5.3}$$

The conservation law we desire is obtained by integrating the Euler equation (4.25) over the cross-space and multiplying by ϵ_{w} to get

$$-\frac{\partial}{\partial t}(\mathscr{L}_\omega) + \frac{\partial}{\partial x_1}(\mathscr{L}_{k_1}) + \frac{\partial}{\partial x_3}(\mathscr{L}_{k_3}) = O(\epsilon_{\mathrm{rf}}, \epsilon_q^3, \epsilon_{\mathrm{w}}^2),$$

or, in view of (5.3),

$$\frac{\partial}{\partial t}(\mathscr{L}_\omega) + \frac{\partial}{\partial x_1}(\Omega_{k_1} \mathscr{L}_\omega) + \frac{\partial}{\partial x_3}(\Omega_{k_3} \mathscr{L}_\omega) = O(\epsilon_{\mathrm{rf}}, \epsilon_q^3, \epsilon_{\mathrm{w}}^2). \tag{5.4}$$

This equation is identical in form to Whitham's law of conservation of wave-action density (cf. Whitham 1974, pp. 393–394). Equation (5.4) identifies Ω_{k_j} as a factor of proportionality between the density and flux of a conserved quantity. It is, therefore, the velocity of propagation of $\mathscr{L}_\omega$.

5.1. *Evolution equations for the wavenumber and wave action density fields*

Several far-reaching results may be derived directly from the definitions of $\boldsymbol{k}$ and ω as derivatives of the phase function $\theta(\boldsymbol{x}, t)$ [cf. (4.27) and (4.28)] and the dispersion relation (4.31). In the following, we will retain the notation convention described in the paragraphs between (4.18) and (4.19) above with regard to the interpretation of the meaning of partial-derivative symbols when both an 'implicit' and an 'explicit' interpretation of partial derivatives of $\Omega(k_1, k_3, x_1, x_3, t)$ with respect to x_1, x_3 and t is possible. Such a convention is necessary to avoid ambiguity since k_1 and k_3 depend on these same variables of differentiation.

A partial differential equation for k_j may be derived by writing the mixed partial derivative of $\theta(x_1, x_3, t)$ with respect to x_j and t in two ways, one with k_j for the 'inner' derivative and the other with $-\omega$ for it. From (4.27) and (4.28) we deduce that

$$\frac{\partial k_j}{\partial t} = -\frac{\partial \omega}{\partial x_j}.$$

Substituting the dispersion relation $\omega = \Omega(k_1, k_3, x_1, x_3, t)$, we get

$$\frac{\partial k_j}{\partial t} = -\frac{\partial}{\partial x_j}[\Omega(k_1, k_3, x_1, x_3, t)], \tag{5.5}$$

which is a nonlinear equation of evolution for the unsteady wavenumber field $k_i(x_1, x_3, t)$. Equation (5.5) may be solved numerically by a step-by-step integration in time subject to the initial conditions

$$k_i(x_1, x_3, 0) = f_i(x_1, x_3) = \text{given}, \quad i \in \{1, 3\}. \tag{5.6}$$

Equation (5.4) for the wave action density may also be written as an equation of evolution. We have, from (5.4),

$$\frac{\partial}{\partial t}(\mathscr{L}_\omega) = -\frac{\partial}{\partial x_1}(\Omega_{k_1}\mathscr{L}_\omega) - \frac{\partial}{\partial x_3}(\Omega_{k_3}\mathscr{L}_\omega) + O(\epsilon_{\mathrm{rf}}, \epsilon_q^3, \epsilon_{\mathrm{w}}^2). \tag{5.7}$$

This equation may also be solved numerically by a step-by-step integration in time, an obvious procedure being to carry out the solution of (5.5) and (5.7) simultaneously subject to the initial conditions (5.6) for k_i and the initial condition

$$\mathscr{L}_\omega(x_1, x_3, 0) = \mathscr{L}_{\omega 0}(x_1, x_3) = \text{given}, \tag{5.8}$$

for $\mathscr{L}_\omega$. One advantage of the above procedure for defining the unsteady evolution of $\mathscr{L}_\omega$ is that the entire calculation can be carried out with real position coordinates (x_1, x_3). This advantage is offset by the disadvantage that further analytical progress is difficult to make and numerical results, if obtained, would be idiosyncratic to a particular set of initial conditions.

5.2. *An alternative approach involving the method of characteristics*

The following notation convention will permit greater economy in the subsequent mathematical development: *if a Greek letter subscript appears twice, it indicates summation over odd indices from* 1 *to* 3. With this convention, (5.5) may be written, with the aid of the chain rule, in the form

$$\frac{\partial k_i}{\partial t} + \Omega_{k_\nu}\frac{\partial k_\nu}{\partial x_i} = -\Omega_{x_i}, \quad i \in \{1, 3\}$$

but

$$\frac{\partial k_\nu}{\partial x_i} = \frac{\partial^2 \theta}{\partial x_i\, \partial x_\nu} = \frac{\partial k_i}{\partial x_\nu}$$

so

$$\frac{\partial k_i}{\partial t} + \Omega_{k_\nu}\frac{\partial k_i}{\partial x_\nu} = -\Omega_{x_i}, \quad i \in \{1, 3\}. \tag{5.9}$$

For general dispersion functions, it is not possible to solve the above system by the method of characteristics in a way that ensures real position coordinates (x_1, x_3) at *all* points of a single characteristic curve. By considering the whole (multiple-parameter) family of characteristic curves, however, it is possible to define within the set of solution points of that family a *physical subset* in which x_1, x_3 and t are all real. Indeed, if the initial-data function on the right-hand side of (5.6) admits an analytic continuation in a neighbourhood of the real axis of each of its two arguments, then the combinations of the variables x_1, x_3, k_1, k_3 and t within the physical subset defined above is reconcilable with the solution of the evolution equation (5.5) for $k_i = k_i(x_1, x_3, t)$ as we will now show. The present approach is similar to the one followed by Itoh (1981) in the case of one-dimensional propagation.

Let (ζ_1, ζ_3) be complex position coordinates. We rewrite (5.9) in the form

$$\frac{\partial k_i}{\partial t} + \Omega_{k_\nu}\frac{\partial k_i}{\partial \zeta_\nu} = -\Omega_{\zeta_i}, \quad i \in \{1, 3\} \tag{5.10}$$

and construct the characteristic curves so that the differential operator on the left-hand side reduces to an ordinary derivative with respect to time. Thus, we require that

$$\frac{\mathrm{d}}{\mathrm{d}t}(\ ) = \frac{\partial}{\partial t}(\ ) + \Omega_{k_\nu}\frac{\partial}{\partial \zeta_\nu}(\ ) \tag{5.11}$$

and hence, that

$$\frac{\mathrm{d}\zeta_i}{\mathrm{d}t} = \Omega_{k_i}, \quad \frac{\mathrm{d}k_i}{\mathrm{d}t} = -\Omega_{\zeta_i}, \tag{5.12a, b}$$

which are complex forms of the so-called *ray equations* of kinematic wave theory (cf. Whitham 1974, §11.5; or Lighthill 1978, §4.5). The corresponding form of the dispersion relation is

$$\omega = \Omega(k_1, k_3, \zeta_1, \zeta_3, t), \tag{5.13}$$

so the system (5.12) forms a closed set. We may write the general solution of (5.12) symbolically in the form

$$\zeta_i = Z_i(t; \zeta_{01}, \zeta_{03}, k_{01}, k_{03}), \quad k_i = K_i(t; \zeta_{01}, \zeta_{03}, k_{01}, k_{03}), \tag{5.14a, b}$$

where the quantities with subscript '0' are the values at $t = 0$ of the corresponding unsubscripted quantities. We suppose that a relationship between k_{0i} and ζ_{0i} exists of the form

$$k_{0i} = f_i(\zeta_{01}, \zeta_{03}), \quad i \in \{1, 3\}, \tag{5.15}$$

where the function on the right-hand side is the analytic continuation of the initial-condition function introduced in (5.6) above. Eliminating k_{01} and k_{03} from (5.14*a, b*) by means of (5.15) induces a functional dependence of the left-hand side quantities upon the three quantities ζ_{01}, ζ_{03} and t, i.e.

$$\zeta_i = \tilde{Z}_i(\zeta_{01}, \zeta_{03}, t), \tag{5.16a}$$

$$k_i = \tilde{K}_i(\zeta_{01}, \zeta_{03}, t), \quad i \in \{1, 3\}. \tag{5.16b}$$

In the above system, there are four equations relating seven quantities and so only three of them are independent. Taking ζ_1, ζ_3 and t to be the independent ones, we may, in principle, express k_i in the form

$$k_i = K_i^{\dagger}(\zeta_1, \zeta_3, t), \quad i \in \{1, 3\} \tag{5.17}$$

Following Itoh (1981), we assign physical meaning only to the subset of points related by (5.17) in which

$$\mathrm{Re}(\zeta_i) = x_i, \quad \mathrm{Im}(\zeta_i) = 0, \quad i \in \{1, 3\} \tag{5.18}$$

so that

$$k_i = K_i^{\dagger}(x_1, x_3, t) \tag{5.19}$$

in this physical subset.

The solution for the wave-action density may be found by the method of characteristics in an analogous way. The complex form of (5.4) is

$$\frac{\partial}{\partial t}(\mathscr{L}_\omega) + \frac{\partial}{\partial \zeta_\nu}(\Omega_{k_\nu}\mathscr{L}_\omega) = O(\epsilon_{\mathrm{rf}}, \epsilon_q^3, \epsilon_{\mathrm{w}}^2).$$

Differentiating the product on the left and substituting (5.11), where possible, we get

$$\frac{\mathrm{d}\mathscr{L}_\omega}{\mathrm{d}t} = -\mathscr{L}_\omega \frac{\partial \Omega_{k_\nu}}{\partial \zeta_\nu} + O(\epsilon_{\mathrm{rf}}, \epsilon_q^3, \epsilon_{\mathrm{w}}^2)$$

$$= -\mathscr{L}_\omega \left(\frac{\partial k_\mu}{\partial \zeta_\nu} \Omega_{k_\nu k_\mu} + \Omega_{k_\nu \zeta_\nu} \right) + O(\epsilon_{\mathrm{rf}}, \epsilon_q^3, \epsilon_{\mathrm{w}}^2), \tag{5.20}$$

where the last equality follows from an application of the chain rule.

One simple approach to the solution of (5.20) is to express the quantity

$$\frac{\partial k_\mu}{\partial \zeta_\nu}$$

in terms of ζ_1, ζ_3 and t by means of (5.17) and then express the quantity

$$\frac{\partial k_\mu}{\partial \zeta_\nu} \Omega_{k_\nu k_\mu} + \Omega_{k_\nu \zeta_\nu}$$

in terms of ζ_{01}, ζ_{03}, k_{01}, k_{03} and t by means of (5.16*a*, *b*). Equation (5.20) could then be written in the form

$$\frac{1}{\mathscr{L}_\omega} \frac{\mathrm{d}\mathscr{L}_\omega}{\mathrm{d}t} = q(t;\, \zeta_{01}, \zeta_{03}, k_{01}, k_{03}) + O(\epsilon_{\mathrm{rf}}, \epsilon_q, \epsilon_{\mathrm{w}}^2),$$

which is of 'variables-separable' type along a characteristic curve.

We will not follow this simple approach for two reasons. First, insights into the development of singularities of $\mathscr{L}_\omega$ along the integration path are difficult to reach by this method. Secondly, there would seem to be practical advantages to a method that calculates the distributions along a characteristic of all the variables one wants in a single integration rather than the sequence of integrations that the above approach requires.

Following Hayes (1970*b*) we will derive equations for the rate of change of $\partial k_i/\partial \zeta_j$ along a characteristic by applying the operator $\partial(\ )/\partial \zeta_j$ to (5.10) and expanding the resulting derivatives by the chain rule. We get

$$\frac{\partial}{\partial t}\left(\frac{\partial k_i}{\partial \zeta_j}\right) + \Omega_{k_\nu} \frac{\partial}{\partial \zeta_\nu}\left(\frac{\partial k_i}{\partial \zeta_j}\right) + \left(\frac{\partial k_\mu}{\partial \zeta_j} \Omega_{k_\mu k_\nu} + \Omega_{\zeta_j k_\nu}\right) \frac{\partial k_i}{\partial \zeta_\nu} = -\left(\frac{\partial k_\mu}{\partial \zeta_j} \Omega_{k_\mu \zeta_i} + \Omega_{\zeta_j \zeta_i}\right).$$

In view of the definition (5.11) of $\mathrm{d}(\ )/\mathrm{d}t$, we may write this equation in matrix notation in the form (cf. Hayes 1970*b*, equation (19))

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{A} = -\boldsymbol{ABA} - \boldsymbol{CA} - \boldsymbol{AC}^{\mathrm{T}} - \boldsymbol{D} \tag{5.21a}$$

where

$$\boldsymbol{A} = \begin{bmatrix} \dfrac{\partial k_1}{\partial \zeta_1} & \dfrac{\partial k_1}{\partial \zeta_3} \\[2ex] \dfrac{\partial k_3}{\partial \zeta_1} & \dfrac{\partial k_3}{\partial \zeta_3} \end{bmatrix}, \quad \boldsymbol{B} = \begin{bmatrix} \Omega_{k_1 k_1} & \Omega_{k_1 k_3} \\[2ex] \Omega_{k_3 k_1} & \Omega_{k_3 k_3} \end{bmatrix}, \tag{5.21b, c}$$

$$\boldsymbol{C} = \begin{bmatrix} \Omega_{k_1 \zeta_1} & \Omega_{k_1 \zeta_3} \\ \Omega_{k_3 \zeta_1} & \Omega_{k_3 \zeta_3} \end{bmatrix}, \quad \boldsymbol{D} = \begin{bmatrix} \Omega_{\zeta_1 \zeta_1} & \Omega_{\zeta_1 \zeta_3} \\ \Omega_{\zeta_3 \zeta_1} & \Omega_{\zeta_3 \zeta_3} \end{bmatrix}, \tag{5.21d, e}$$

and where the symbol $[\ ]^{\mathrm{T}}$ denotes the transpose of the matrix operand. In the same notation, (5.20) takes the form

$$\frac{\mathrm{d}\mathscr{L}_\omega}{\mathrm{d}t} = -\mathscr{L}_\omega \operatorname{tr}(\boldsymbol{AB}+\boldsymbol{C}) + O(\epsilon_{\mathrm{rf}}, \epsilon_q^3, \epsilon_{\mathrm{w}}^2) \tag{5.22}$$

where the symbol tr( ) denotes the trace of the matrix operand.

The initial condition for the matrix $\boldsymbol{A}$ must be consistent with (5.15) and (5.21b). We have, therefore,

$$\boldsymbol{A}_0 = \boldsymbol{A}_0(\zeta_{01}, \zeta_{03}) = \begin{bmatrix} \dfrac{\partial f_1}{\partial \zeta_{01}} & \dfrac{\partial f_1}{\partial \zeta_{03}} \\ \dfrac{\partial f_3}{\partial \zeta_{01}} & \dfrac{\partial f_3}{\partial \zeta_{03}} \end{bmatrix}. \tag{5.23}$$

We suppose that the initial condition (5.8) for $\mathscr{L}_\omega$ has an analytic continuation of the form

$$\mathscr{L}_\omega(\zeta_1, \zeta_3, 0) = \mathscr{L}_{\omega 0}(\zeta_{01}, \zeta_{03}) = \text{given}. \tag{5.24}$$

Now the 2×2 matrix $\boldsymbol{A}$ is symmetric [as follows from the equation immediately above (5.9)]. The matrix equation (5.21a) is equivalent to three independent scalar equations. The system consisting of the four scalar equations (5.12a,b), the three scalar equations (5.21), and the scalar equation (5.22) is, therefore, a system of order eight for the eight unknowns

$$\zeta_1, \zeta_3, k_1, k_3, \frac{\partial k_1}{\partial \zeta_1}, \frac{\partial k_1}{\partial \zeta_3} = \frac{\partial k_3}{\partial \zeta_1}, \frac{\partial k_3}{\partial \zeta_3}, \mathscr{L}_\omega.$$

The general solution for the first four quantities is given by (5.14a,b). For the remaining four, we may write

$$\frac{\partial k_1}{\partial \zeta_1} = K_{11}(t;\, \zeta_{0i}, k_{0i}, \boldsymbol{A}_0), \tag{5.25a}$$

$$\frac{\partial k_1}{\partial \zeta_3} = \frac{\partial k_3}{\partial \zeta_1} = K_{13}(t;\, \zeta_{0i}, k_{0i}, \boldsymbol{A}_0), \tag{5.25b}$$

$$\frac{\partial k_3}{\partial \zeta_3} = K_{33}(t;\, \zeta_{0i}, k_{0i}, \boldsymbol{A}_0), \tag{5.25c}$$

$$\mathscr{L}_\omega = \mathscr{L}_\omega^{\ddagger}(t;\, \zeta_{0i}, k_{0i}, \boldsymbol{A}_0, \mathscr{L}_{\omega 0}). \tag{5.25d}$$

Substitution of the initial conditions (5.15), (5.23) and (5.24) to eliminate k_{0i}, $\boldsymbol{A}_0$ and $\mathscr{L}_{\omega 0}$ from the right-hand sides induces a functional dependence of the form

$$\frac{\partial k_i}{\partial \zeta_j} = \tilde{K}_{ij}(\zeta_{01}, \zeta_{03}, t), \quad \mathscr{L}_\omega = \tilde{\mathscr{L}}_\omega(\zeta_{01}, \zeta_{03}, t). \tag{5.26a,b}$$

The system (5.16a,b), (5.26a,b) is equivalent to eight scalar equations relating eleven scalar quantities and so only three of these quantities are independent. Taking ζ_1, ζ_3 and t to be the independent ones [as we did before, in the paragraph between (5.16) and (5.17)], we may write

$$\frac{\partial k_i}{\partial \zeta_j} = K_{ij}^{\dagger}(\zeta_1, \zeta_3, t), \quad \mathscr{L}_\omega = \mathscr{L}_\omega^{\dagger}(\zeta_1, \zeta_3, t). \tag{5.27a,b}$$

The physically meaningful subset of points related by these equations is the subset defined by (5.18), i.e.

$$\frac{\partial k_i}{\partial x_j} = K^{\dagger}_{ij}(x_1, x_3, t), \quad \mathscr{L}_\omega = \mathscr{L}^{\dagger}_\omega(x_1, x_3, t). \tag{5.28a, b}$$

The above solution by the method of characteristics is predicated on the availability of analytic continuations of all the initial-data functions and of the dispersion function for complex horizontal position coordinates ζ_1 and ζ_3. The process would fail if, for example, any of the requisite analytic continuations did not exist. The requirement that such analytic continuations exist should not pose a serious restriction on the usefulness of the method if the complex characteristics satisfy a condition of the form

$$|\mathrm{Im}\,(\zeta_i)| \ll |\mathrm{Re}\,(\zeta_i)|, \quad i \in \{1, 3\} \tag{5.29}$$

for the entire range of t over which solutions are sought. In the next subsection, we will discuss a class of special cases in which the condition (5.29) holds. We will show that within this class the solution for $\mathscr{L}_\omega$ may exhibit a singularity along a characteristic for finite time.

5.3. *Approximate solutions when Ω_{ζ_i} and Ω_t are both small*

A set of conditions sufficient to ensure that (5.29) is satisfied within a time interval $t \in (0, T)$ are

(i) $$|\Omega_{\zeta_i}| = O(\epsilon_{\mathrm{rf}}), \quad |\Omega_t| = O(\epsilon_{\mathrm{rf}}) \quad \text{for all } t \in (0, T) \tag{5.30a}$$

(ii) $$|\mathrm{Im}\,(\Omega_{k_i})| \ll |\mathrm{Re}\,(\Omega_{k_i})|, \quad |\mathrm{Im}\,(\zeta_i)| \ll |\mathrm{Re}\,(\zeta_i)| \quad \text{for a single } t = t_1 \in (0, T). \tag{5.30b}$$

When these conditions are satisfied, the ray equations (5.12*a*, *b*) reduce to

$$\frac{\mathrm{d}\zeta_i}{\mathrm{d}t} = \Omega_{k_i}, \quad \frac{\mathrm{d}k_i}{\mathrm{d}t} = O(\epsilon_{\mathrm{rf}}), \tag{5.31a, b}$$

showing that k_i changes very little along a characteristic curve. Now Ω_{k_i} depends on the variables k_i, ζ_i and t and the condition (5.30*a*) ensures that the dependencies upon ζ_i and t are weak. It follows that the right-hand side of (5.31*a*) is virtually constant. Equations (5.30*b*) then ensure that (5.29) is satisfied, as stated.

Now, ϵ_{rf} is a measure of the weakness of the dependency of background flow quantities such as U_1 and U_3 upon the variables x_1, x_3 and t. Also, the dependence of the dispersion function $\omega = \Omega(k_1, k_3, x_1, x_3, t)$ upon these variables is due solely to the appearance of $\boldsymbol{U}$ as a coefficient in the Rayleigh stability equation [cf. the equation between (4.28) and (4.29)]. Since we have already employed the assumption $\epsilon_{\mathrm{rf}} \ll 1$ in all the derivations prior to those of the preceding section, we see that (5.30*a*) will be satisfied under approximations already stated, provided that $T \leqq l/Q$ [cf. (2.3) and (2.4)].

The propagation equations for $\boldsymbol{A}$ and $\mathscr{L}_\omega$ simplify under conditions (5.30*b*). We note that all the elements of the matrix $\boldsymbol{C}$ in (5.21*d*) are $O(\epsilon_{\mathrm{rf}})$ and all those of matrix $\boldsymbol{D}$ in (5.21*e*) are $O(\epsilon^2_{\mathrm{rf}})$. It follows that (5.21*a*) and (5.22) take the simpler forms

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{A} = -\boldsymbol{ABA} + O(\epsilon_{\mathrm{rf}}), \tag{5.32}$$

and

$$\frac{\mathrm{d}\mathscr{L}_\omega}{\mathrm{d}t} = -\mathscr{L}_\omega\,\mathrm{tr}\,(\boldsymbol{AB}) + O(\epsilon_{\mathrm{rf}}, \epsilon^3_q, \epsilon^2_{\mathrm{w}}). \tag{5.33}$$

Since $\boldsymbol{A}$ and $\boldsymbol{B}$ are symmetric, we have

$$\operatorname{tr}(\boldsymbol{AB}) = \operatorname{tr}\{(\boldsymbol{AB})^{\mathrm{T}}\} = \operatorname{tr}\{\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}\} = \operatorname{tr}(\boldsymbol{BA}),$$

so (5.33) can be written

$$\frac{1}{\mathscr{L}_\omega}\frac{\mathrm{d}\mathscr{L}_\omega}{\mathrm{d}t} = -\operatorname{tr}(\boldsymbol{BA}) + O(\epsilon_{\mathrm{rf}}, \epsilon_q, \epsilon_{\mathrm{w}}^2). \tag{5.34}$$

Noting the identity

$$\frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{A}^{-1}\boldsymbol{A}) = \frac{\mathrm{d}\boldsymbol{A}^{-1}}{\mathrm{d}t}\boldsymbol{A} + \boldsymbol{A}^{-1}\frac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t} = [0], \tag{5.35}$$

we have

$$\frac{\mathrm{d}\boldsymbol{A}^{-1}}{\mathrm{d}t} = -\boldsymbol{A}^{-1}\frac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t}\boldsymbol{A}^{-1}.$$

Substituting (5.32) into the right-hand side to eliminate $\mathrm{d}\boldsymbol{A}/\mathrm{d}t$, we get

$$\frac{\mathrm{d}\boldsymbol{A}^{-1}}{\mathrm{d}t} = \boldsymbol{B} + O(\epsilon_{\mathrm{rf}}). \tag{5.36}$$

Substituting this result into (5.34) to eliminate $\boldsymbol{B}$, we get

$$\left.\begin{aligned} \frac{1}{\mathscr{L}_\omega}\frac{\mathrm{d}\mathscr{L}_\omega}{\mathrm{d}t} &= -\operatorname{tr}\left(\frac{\mathrm{d}\boldsymbol{A}^{-1}}{\mathrm{d}t}\boldsymbol{A}\right) + O(\epsilon_{\mathrm{rf}}, \epsilon_q, \epsilon_\omega^2), \\ &= \operatorname{tr}\left(\boldsymbol{A}^{-1}\frac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t}\right) + O(\epsilon_{\mathrm{rf}}, \epsilon_q, \epsilon_{\mathrm{w}}^2), \end{aligned}\right\} \tag{5.37}$$

where the last equality follows from another application of (5.35).

Hayes (1970*b*) has pointed out that the first term on the right-hand side of (5.37) is equal to the logarithmic derivative of the determinant of $\boldsymbol{A}$ with respect to t. To show this, we let $\det\boldsymbol{A}$ denote the determinant of $\boldsymbol{A}$ and let a_{ij} denote the element in the ith row and the jth column of $\boldsymbol{A}$. Then $\det\boldsymbol{A}$ is a function of the various a_{ij}. By application of the chain rule, we have

$$\frac{\mathrm{d}}{\mathrm{d}t}(\det\boldsymbol{A}) = \sum_{i,j}\frac{\partial\det\boldsymbol{A}}{\partial a_{ij}}\frac{\mathrm{d}a_{ij}}{\mathrm{d}t},$$

where the sum is over all the elements of $\boldsymbol{A}$. But

$$\frac{\partial\det\boldsymbol{A}}{\partial a_{ij}} \equiv A_{ij}$$

is the cofactor of the element a_{ij} in $\boldsymbol{A}$ and, by Cramer's rule, $\boldsymbol{A}_{ij}/\det\boldsymbol{A}$ is the element in the jth row and the ith column of $\boldsymbol{A}^{-1}$. It follows that

$$\frac{1}{\det\boldsymbol{A}}\frac{\mathrm{d}}{\mathrm{d}t}(\det\boldsymbol{A}) = \operatorname{tr}\left(\boldsymbol{A}^{-1}\frac{\mathrm{d}\boldsymbol{A}}{\mathrm{d}t}\right),$$

as stated, and (5.37) takes the integrable form

$$\frac{1}{\mathscr{L}_\omega}\frac{\mathrm{d}\mathscr{L}_\omega}{\mathrm{d}t} = \frac{1}{\det\boldsymbol{A}}\frac{\mathrm{d}}{\mathrm{d}t}(\det\boldsymbol{A}) + O(\epsilon_{\mathrm{rf}}, \epsilon_q, \epsilon_{\mathrm{w}}^2) \tag{5.38}$$

whence

$$\frac{\mathscr{L}_\omega}{\mathscr{L}_{\omega 0}} = \frac{\det\boldsymbol{A}}{\det\boldsymbol{A}_0}[1 + O(\epsilon_{\mathrm{rf}}, \epsilon_q, \epsilon_{\mathrm{w}}^2)].$$

Now the function $\boldsymbol{B}$ in (5.36) depends only on the solutions of the ray equations (5.31 a, b) and the time and is independent of the elements of $\boldsymbol{A}$. It follows from (5.36) that

$$\boldsymbol{A} = \left(\boldsymbol{A}_0^{-1} + \int_0^t \boldsymbol{B}\, d\tau\right)^{-1} [1 + O(\epsilon_{\mathrm{rf}})],$$

and (5.38) takes the form

$$\frac{\mathscr{L}_\omega}{\mathscr{L}_{\omega 0}} = \frac{\det \boldsymbol{A}_0^{-1}}{\det\left(\boldsymbol{A}_0^{-1} + \int_0^t \boldsymbol{B}\, d\tau\right)} [1 + O(\epsilon_{\mathrm{rf}})]. \tag{5.39}$$

This approximate solution invites comparison with certain well-known results regarding the far-field dispersed wave pattern in a uniform medium. If $\tilde{\varphi}_0(\boldsymbol{k})$ is the Fourier transform of the initial distribution of some dependent variable $\varphi(\boldsymbol{x}, t)$, then the method of stationary phase may be employed to derive the result (cf. Whitham 1974, equation (11.41))

$$\varphi(\boldsymbol{x}, t) \sim \tilde{\varphi}_0(\boldsymbol{k}) \left(\frac{2\pi}{t}\right)^{\frac{1}{2}n} (\det \boldsymbol{B})^{-\frac{1}{2}} \exp[i(\boldsymbol{k}\cdot\boldsymbol{x} - \Omega(\boldsymbol{k})t + \zeta)],$$

where $x_i/t = \Omega_{k_i}$, n = number of dimensions in the propagation space, ζ = a real constant determined by the number of path rotations employed in the stationary-phase algorithm. In the case $n = 2$, we find

$$\varphi\varphi^* \sim \tilde{\varphi}_0 \tilde{\varphi}_0^* \, 4\pi^2 \frac{\exp[-2\,\mathrm{Im}\{\boldsymbol{k}\cdot\boldsymbol{x} - \Omega(\boldsymbol{k})t\}]}{\det(\boldsymbol{B}t)}. \tag{5.40}$$

Comparing the form taken by (5.39) in the limit $t \to \infty$ for a uniform steady medium with the above result, we see that the quantity $\mathscr{L}_\omega/\mathscr{L}_{\omega 0}$ represents the factor in the complete expression for the square amplitude associated with amplitude changes by dispersion, i.e. $\mathscr{L}_\omega$ does not display any effects due to exponential amplification or decay along the ray. This feature of $\mathscr{L}_\omega$ is expected from its definition as a bilinear product of the 'script' and the 'italic' trial solutions in (4.12 a–d). If either of the sets of variables (b_1, p) or $(\mathscr{b}_1, \mathscr{p})$ represents an amplified mode, then the other will represent a damped one. The use of a bilinear Lagrangian density causes their respective exponential amplification rates to cancel. We will return in §6 to the question of how to reinsert the appropriate exponential amplification factor to get an expression for the 'observed' square amplitude of an amplified mode.

The well-known far-field 'caustics' are the curves along which the determinant in the denominator of (5.40) vanishes. The equations of kinematic wave theory are not valid on the caustics themselves but do properly indicate their location. Careful analysis involving the use of the Airy function (cf. Lighthill 1978, §4.11) show that the actual square amplitude on the caustic is finite, as expected, the main effect of the caustic being to change the negative power of t in the formula for $\varphi\varphi^*$ from -2 to some other negative number.

Equation (5.32) may be regarded as a simple example of a *matrix Riccati equation*. Hille (1969, appendix C.3) presents an interesting discussion of matrix Riccati equations in general and establishes several analogies between the kinds of singularities that occur in the matrix Ricatti equation and those that occur in the more familiar scalar Riccati equation. The concept of a *movable singular point* plays an important role in the theory. A singularity in the solution of a nonlinear ordinary differential equation is called 'movable' if its location in the space of the independent variable is dependent upon the initial conditions and can be placed anywhere in that space by suitable choice of those initial conditions. If, for example, $a = a(t)$ is a solution of the scalar equation

$$\frac{\mathrm{d}a}{\mathrm{d}t} = -ba^2,$$

(where b is a constant), then the solution for $a(t)$, subject to the initial condition

$$a = a_0 \quad \text{at } t = 0$$

is

$$a = (a_0^{-1} + bt)^{-1},$$

which is precisely analogous to the matrix equation between (5.38) and (5.39). The function $a(t)$ has a pole at $t = -(ba_0)^{-1}$ whose location can be placed anywhere in the interval $(-\infty, \infty)$ by appropriate choice of a_0.

Movable singular points in the solution for the matrix $\boldsymbol{A}$ and, hence, in the solution for $\mathscr{L}_\omega$, will exist whenever the determinant in the denominator of (5.39) vanishes. Such movable singular points would no more represent infinitely large values of real physical quantities than the far-field caustics do. As was the case with the far-field caustics, however, kinematic wave theory provides a clue to the locations of interesting phenomena, particularly those characterized by untypically large amplitudes. If we approximate the matrix $\boldsymbol{B}$ by its initial value $\boldsymbol{B}_0$ along a ray and introduce the notation $\lambda = -1/t$, then (5.39) may be written

$$\frac{\mathscr{L}_\omega}{\mathscr{L}_{\omega 0}} = \frac{\lambda^2}{\det(\boldsymbol{A}_0 \boldsymbol{B}_0 - \lambda \boldsymbol{I})} [1 + O(\epsilon_{\mathrm{rf}}, \epsilon_q, \epsilon_{\mathrm{w}}^2)].$$

If the matrix product $\boldsymbol{A}_0 \boldsymbol{B}_0$ [whose elements are approximately the elements of the group velocity-gradient tensor $(\partial \Omega_{k_i}/\partial x_i)_0$ if $\epsilon_{\mathrm{rf}} \ll 1$] has a negative real eigenvalue for a particular ray, then that eigenvalue is minus the reciprocal of the time when that ray would pass into a near-field movable singular point.

Such near-field focii associated with non-uniformity of the initial conditions must be well known to people who frequently apply kinematic wave theory. The author is unaware of any discussion of them in the context of shear-flow instability waves, however. Theories of laminar–turbulent transition and of turbulence maintenance in fully developed turbulent flows might well profit from consideration of such singularities.

5.4. *On steady solutions of the equations for k_i and $\mathscr{L}_\omega$*

A third type of focus, studied in detail by Landahl (1972) and Itoh (1981) deserved mention in this context. If a reference frame can be found in which the background flow velocity profiles are stationary in time, then steady-flow solutions of the equations for k_i and $\mathscr{L}_\omega$ satisfying

$$\frac{\partial k_i}{\partial t} = 0, \quad \frac{\partial \mathscr{L}_\omega}{\partial t} = 0, \tag{5.41 a, b}$$

may be possible. The reference frame in which the above equations hold may, of course, be moving relative to the fluid particles at the wall elevation. In the case where the above two equations hold, specification of initial data in the form (5.6), (5.8) is inappropriate. Rather, initial data should be specified along an 'inflow-boundary' contour Γ described, say, by an equation of the form

$$x_1 = \text{constant on } \Gamma.$$

If (5.41 a) is substituted into the equation immediately before the wavenumber evolution equation (5.5), we get $\partial\omega/\partial x_j = 0$, which implies that ω is a function of time only. On the other hand, if the background-flow velocity profiles are stationary in time, then the dispersion relation reduces to

$$\omega = \Omega(k_1, k_3, x_1, x_3)$$

from which it follows that $\omega = \omega_0$ is independent of time as well. We obtain, therefore,

$$\omega_0 = \Omega(k_1, k_3, x_1, x_3). \tag{5.42}$$

For any given point on the inflow boundary contour Γ, the values of x_1 and x_3 are specified, so equation (5.42) defines a relationship that must hold between the boundary values of the two quantities k_1 and k_3. Thus, we refrain from specifying *both* of the quantities k_1 and k_3 independently on Γ. We specify instead a second relationship between k_1 and k_3 of the form

$$f_0(k_1, k_3, x_3) = 0 \quad \text{on } \Gamma, \tag{5.43}$$

which, together with (5.42), determines the boundary values of k_1 and k_3. The equations for the steady wavenumber field $k_i(x_1, x_3)$ may be found by substituting (5.41 a) into (5.9), or, equivalently, by applying the operator $\partial(\ )/\partial x_j$ to (5.42). The result is the system of equations

$$\Omega_{k_\nu} \frac{\partial k_i}{\partial x_\nu} = -\Omega_{x_i}, \quad i \in \{1, 3\}, \tag{5.44}$$

which may be solved, in principle, by a step-by-step integration in x_i. According to (5.41 b), the wave-action-density conservation equation (5.4) takes the form

$$\frac{\partial}{\partial x_1}(\Omega_{k_1}\mathscr{L}_\omega) + \frac{\partial}{\partial x_3}(\Omega_{k_3}\mathscr{L}_\omega) = O(\epsilon_{\mathrm{rf}}, \epsilon_q^3, \epsilon_{\mathrm{w}}^2). \tag{5.45}$$

We suppose further that initial conditions of the form

$$\mathscr{L}_\omega = g_0(x_3) \quad \text{on } \Gamma \tag{5.46}$$

have been specified. For the remainder of this subsection, we restrict attention to the case in which the system (5.42)–(5.46) has a unique solution for the functions $k_i(x_1, x_3)$ and $\mathscr{L}_\omega(x_1, x_3)$ in some non-trivial region of the (x_1, x_3)-plane neighbouring the initial contour Γ.

If the dispersion relation and the initial-data functions are both independent of x_3, then the problem is two-dimensional and the equation

$$\frac{\partial}{\partial x_3}(\Omega_{k_3}\mathscr{L}_\omega) = O(\epsilon_{\mathrm{rf}}, \epsilon_q^3, \epsilon_{\mathrm{w}}^2),$$

holds identically for all x_3. This equation states that the flux of wave action density in the x_3 direction is independent of the x_3 coordinate. It is possible to define a class of three-dimensional flows in which the above equation holds for one or more 'spanwise stations' $x_3 = b =$ constant. We expect, for example, that such spanwise stations will exist if the flow is periodic in the x_3 direction or if the flow has a longitudinal plane of symmetry. This class of three-dimensional flows is the one considered by Landahl (1972). On the line $x_3 = b$, it follows from (5.45) that

$$\mathscr{L}_\omega \propto \frac{1}{\Omega_{k_1}} \quad \text{on } x_3 = b,$$

so $\mathscr{L}_\omega$ becomes singular at a point where the denominator on the right vanishes. This is Landahl's 'breakdown' criterion expressed in a reference frame where the flow is stationary. Landahl (1972) argued that the onset of violent small-scale secondary motions in the experiments of Klebanoff, Tidstrom & Sargent (1962) could be accounted for by application of the above criterion. In his original work on near-field focusing and in a subsequent analysis (Landahl 1982) of the dynamics of wavetrains and packets in both the near and far fields, Landahl has restricted attention to the case when the inequality

$$|\mathrm{Im}\,(\Omega_{k_i})| \ll |\mathrm{Re}\,(\Omega_{k_i})|$$

holds. The present derivation recovers Landahl's breakdown criterion without imposing this restriction.

6. Discussion

6.1. *'Observed' square amplitude*

In the preceding section, we have tacitly regarded $\mathscr{L}_\omega$ as a square-amplitude measure. As the discussion in the paragraph following (5.40) indicates, however, the observed square amplitude of an amplified wavetrain would contain $\mathscr{L}_\omega$ as a significant factor but not the only factor, since $\mathscr{L}_\omega$ does not contain information about exponential growth or decay. Fortunately, the exponential growth factor is very simple to reinsert. By inspection of the trial solutions (4.12*a–d*), we see immediately that if either of the sets of variables (b_i, p) or $(\not{b}_i, \not{p})$ represents an amplified wave, then the other represents a damped one. Suppose, for definiteness, that the set (b_i, p) represents the amplified wave. Then, letting

$$\theta = \mathrm{Re}\,\{\theta\} + \mathrm{i}\,\mathrm{Im}\,\{\theta\}$$

we see that the exponential amplification factor in the variables (b_i, p) will be

$$\exp\,[-\mathrm{Im}\,\{\theta\}].$$

A natural measure for the observed square amplitude A^2 of an amplified wave which allows for both exponential amplification and amplitude changes due to dispersion and focusing is

$$A^2 = \mathscr{L}_\omega \exp\,[-2\,\mathrm{Im}\,\{\theta\}]. \tag{6.1}$$

Eliminating $\mathscr{L}_\omega$ between (6.1) and (5.4), we obtain a 'non-conservation' law for A^2 of the form

$$\frac{\partial}{\partial t}(A^2) + \frac{\partial}{\partial x_1}(\Omega_{k_1} A^2) + \frac{\partial}{\partial x_3}(\Omega_{k_3} A^2)$$
$$= 2A^2(\mathrm{Im}\,\{\Omega\} - \Omega_{k_1}\,\mathrm{Im}\,\{k_1\} - \Omega_{k_3}\,\mathrm{Im}\,\{k_3\}) + O(\epsilon_{\mathrm{rf}}, \epsilon_q^3, \epsilon_{\mathrm{w}}^2), \tag{6.2}$$

which is similar to equation (32) of Landahl (1982). We remark that this equation is also satisfied by the quantity $\varphi\varphi^*$ in (5.40), i.e. the square-amplitude measure normally considered in analyses of the far-field wave pattern by the method of stationary phase.

The decomposition of the square amplitude A^2 into a part $\mathscr{L}_\omega$ associated with dispersion and focusing, and another part $\exp(-2\theta_i)$ associated with exponential amplification is convenient. There is a deeper significance of this decomposition which deserves mention, however. We employed the bilinear variational principle

(4.2) in combination with the trial solutions (4.12a–d) in order to ensure that the exponential amplification factors of the trial solutions would cancel upon substitution into the bilinear Lagrangian density, which then depended only on the derivatives of θ and not on the undifferentiated form of θ. Each independently variable quantity in the Lagrangian density function will, in the application of the variational calculus, yield a corresponding Euler equation. Variables in the Lagrangian density which appear only as derivatives give rise to Euler equations in the form of conservation laws, and conversely, as pointed out in the book by Goldstein (1980, §12.4) who refers to such variables as 'cyclic'.

If, instead of applying the bilinear variational principle (4.2), we had attempted to apply the 'quadratic' principle (4.1) and had substituted into it only the amplified trial solutions (4.12a,b), we would have encountered two difficulties. First, the phase-averaged Lagrangian density would contain the factor $\exp(-2\theta_i)$ so that θ would be 'non-cyclic' and would not give rise to an Euler equation in the form of a conservation law. Second, it is not clear how to interpret the derivative of the real quantity $\exp(-2\theta_i)$ with respect to the complex quantity θ which one must employ in the derivation of the Euler equation corresponding to independent variations of θ.

In summary, the use of a bilinear variational principle with trial functions of opposite exponential amplification properties seems to be necessary if the variational formalism is to be unambiguous and, in particular, if the resulting Euler equation for $\mathscr{L}_\omega$ is to have the form of a conservation law. Such a conservation law can be derived even if the observed square amplitude A^2 obeys a non-conservation law such as (6.2).

6.2. *Comparison between the results of the present theory and those of other investigators*

The present work is an analysis of the dependency of amplitude, frequency, wavenumber and other parameters of a train of shear-flow instability waves upon the initial distributions of those parameters and upon slow variations of the background medium through which they propagate. Viewed in these terms, the problem addressed in this work is comparable with the problems addressed by Nayfeh (1980) and Itoh (1981).

Though there is considerable overlap between the sets of assumptions employed in these three works, no one of them addresses a problem that is, strictly speaking, a special case of any other. In Itoh's problem, the analysis is restricted to two-dimensional flow. Slow streamwise variations and a limited kind of time dependency of the background flow are taken into account, however, as are the effects of viscosity. The time dependency of the background flow is limited to the kind that can be removed by an appropriate choice of translating reference frame, leaving a steady-flow problem of the sort considered in §5.4 above.

In Nayfeh's problem, the background flow is assumed to be steady in time in a reference frame fixed to the wall. Nayfeh's analysis allows for independent functional dependencies of the horizontal velocity components U_1 and U_2 upon the cross-stream coordinate x_2, a complication that we will henceforth refer to as 'skewness' of the background-flow profiles. Slow variations of U_1 and U_3 with respect to the horizontal space coordinates are also allowed for, as are the effects of viscosity. The assumption of strict steadiness of the background flow leads Nayfeh to attribute horizontal non-uniformities of the background flow [cf. the last sentence before equation (8) of

Nayfeh 1980] to the effects of boundary-layer growth and thus to equate the small parameter ϵ that measures horizontal non-uniformity of the background flow to the reciprocal of the Reynolds number R. In this way, the steady-background-flow assumption implies that the background flow becomes progressively 'flatter' as R tends to infinity.

In the present work, skewness and slow horizontal non-uniformities of the background-flow profiles are allowed for, as in Nayfeh's problem, but viscosity is ignored. Thus, in contrast to Nayfeh's problem, horizontal non-uniformities of the background flow may *not* be attributed to finite-Reynolds-number effects. They must, instead, be attributed to a *temporally evolving* larger-scale structure in the background flow.

All three theories reduce to the same problem when one restricts attention to two-dimensional inviscid flow and ignores streamwise variations of the background flow. Even under such an idealization, one is free to prescribe spatially non-uniform initial distributions of the parameters of the disturbance wavetrain and to deduce from the law of conservation of wave action density the possibility of far-field caustics and near-field movable singularities of the sort discussed in §5.3 above.

Equation (5.4) above is our mathematical statement of the law of conservation of (bilinear) wave-action density for shear-flow instability waves and is of exactly the same form as equation (140) of Nayfeh (1980). In an analysis of the instability of plane Poiseuille flow, Itoh (1980) derived the same conservation law (cf. equation 2.18 of that paper). In his later work, however, Itoh chose to write his amplitude-propagation law in a different form, namely as a formula for the rate of change of the *first* power of the amplitude [cf. Itoh (1981), equation 2.13]. The question naturally arises as to whether the various second-order amplitude measures, whose conservation laws are derived by Nayfeh, Itoh and in the present work, are mathematically equivalent.

The answer is no, for reasons that we will elaborate on presently. Given that not all second-order amplitude measures are the same, the next question that arises is whether more than one second-order amplitude measure can represent a conserved quantity. The answer to the latter question is yes and the variational formalism provides a very clear indication of why.

If, in Nayfeh's equations, we substitute $\epsilon = 1/R$ and set $R = \infty$, then the background-flow velocity profiles take the form

$$\left.\begin{aligned} U_i &= U_i(x_2), \quad i \in \{1,3\} \\ U_2 &= 0, \end{aligned}\right\} \tag{6.3}$$

and the small-disturbance equations of motion [Nayfeh's equations (13)–(16)] take the form

$$\operatorname{lsd}(u_i) + u_2\left(\frac{\partial U_1}{\partial x_2}\,\delta_{1i} + \frac{\partial U_3}{\partial x_2}\,\delta_{3i}\right) + \frac{\partial}{\partial x_i}(p) = 0, \tag{6.4a}$$

$$\frac{\partial u_j}{\partial x_j} = 0, \tag{6.4b}$$

where the density is absent owing to Nayfeh's use of non-dimensional variables. These equations are equivalent to (2.9a, b) above under analogous assumptions regarding horizontal uniformity of the background flow. Nayfeh's wave action density is a bilinear expression involving solutions of the above system as one factor and solutions of an adjoint system as the other. Substituting $\epsilon = 1/R = 0$ into Nayfeh's equations (129)–(132), the adjoint system becomes

$$-\operatorname{lsd}(u_i^*) + \left(u_1^* \frac{\partial U_1}{\partial x_2} + u_3^* \frac{\partial U_3}{\partial x_2}\right)\delta_{2i} - \frac{\partial (p^*)}{\partial x_i} = 0, \tag{6.5a}$$

$$-\frac{\partial u_j^*}{\partial x_j} = 0, \tag{6.5b}$$

where the star denotes the adjoint system of variables. Introducing travelling-wave trial solutions of the form

$$p = A\tilde{p}\,\mathrm{e}^{\mathrm{i}\theta}, \tag{6.6a}$$

$$u_i = A\tilde{u}_i\,\mathrm{e}^{\mathrm{i}\theta} \quad i \in \{1,2,3\}, \tag{6.6b}$$

$$p^* = A^*\tilde{p}^*\,\mathrm{e}^{\mathrm{i}\theta^*} \tag{6.6c}$$

$$u_i^* = A^*\tilde{u}_i^*\,\mathrm{e}^{\mathrm{i}\theta^*}, \quad i \in \{1,2,3\}, \tag{6.6d}$$

and defining the wavenumber components and frequency by

$$-\frac{\partial \theta}{\partial t} = \omega = \frac{\partial \theta^*}{\partial t}, \tag{6.7a}$$

$$\frac{\partial \theta}{\partial x_i} = k_i = -\frac{\partial \theta^*}{\partial x_i}, \quad i \in \{1,3\} \tag{6.7b}$$

Nayfeh arrives at a wave-action-density conservation law similar to (5.4) above but with the quantity

$$AA^* \int_0^\infty \tilde{u}_j \tilde{u}_j^* \,\mathrm{d}x_2 \tag{6.8}$$

in place of $\mathscr{L}_\omega$.

The definitions of frequency and wavenumber that apply to the 'starred' and 'non-starred' variables of Nayfeh are completely consistent with those that apply to the 'script' and 'italic' variables in the present work.

The corresponding formula that relates wave-action density in the present work to an integral over the mode shape functions may be derived from (4.16), (5.1), and the notational definitions (3.8), (4.27) and (4.28). Setting $h = \infty$ in (5.1), we obtain

$$\mathscr{L}_\omega = \int_0^\infty 2(\omega - k_1 U_1 - k_3 U_3)(\tilde{b}_j \tilde{\ell}_j)\,\mathrm{d}x_2. \tag{6.9}$$

The quantity $\rho\mathscr{L}_\omega$ has the dimensions of energy per unit of horizontal area divided by frequency. Although Nayfeh's equations are written in non-dimensional form, one may show that ρ times the dimensional counterpart of the quantity (6.8) would have the dimensions of energy per unit of horizontal area. On dimensional grounds, therefore, Nayfeh's wave-action density is not the same as ours.

From the standpoint of mathematical rigor, there is nothing wrong with Nayfeh's derivation of the conservation equation satisfied by the quantity (6.8). Indeed, we will show below how the same conservation law can be derived by a variational method. From the standpoint of *physical interpretability*, however, the need to employ solutions of the adjoint system of equations (which have no apparent physical meaning) as a factor in the resulting expression for the wave action density can only be viewed as a disadvantage.

It often happens that a system of equations that is not self-adjoint may be transformed to an equivalent system which is self-adjoint by a simple change of variables. Such is the case in the present work where the change of variable (3.11) transforms the system (2.9a, b) (which is not self-adjoint) to the system (3.10a, b) (which is). The system of equations (4.3a, b) satisfied by the variables ($\mathscr{b}_t, \mathscr{p}$) is identical with the system (3.10a, b) satisfied by (b_t, p). Neither system of equations is any more difficult to interpret physically than the other, both being transformed versions of the small-disturbance equations of motion. In this sense, the definition of wave action density employed herein is less vulnerable to criticism on the grounds of physical opacity than is Nayfeh's definition.

It is appropriate to note, however, that the self-adjointness of the small-disturbance equations of motion when written in terms of the variables (b_t, p) disappears when viscous effects are included. One of the main goals of this work, was to derive the law of conservation of wave action density in shear waves (which are non-conservative even when viscous effects are ignored) in a form which involves physically meaningful quantities for all the factors in $\mathscr{L}_\omega$. The difficulty of accomplishing this task in the viscous case is what motivated us to restrict attention to the inviscid problem.

We now return to the question of whether different second-order amplitude measures can satisfy similar conservation equations, specifically equations of the form (5.4). One of the most striking features of the derivations of the equation of conservation of wave action density by Hayes (1970a) and Whitham (1974) in the case of conservative systems is the ability of those authors to derive the equation from a *generic class* of variational principles rather than one which is specific to a particular set of partial differential equations governing one problem. The only conditions that must be satisfied by a variational principle to ensure that it leads to the law of conservation of wave-action density are: first, that the original partial differential equations of motion of the system admit solutions in the form of slowly varying wavetrains with a phase function $\theta(x, t)$; secondly, that substitution of trial solutions in the form of slowly varying wavetrains into the variational principle yields, upon averaging over one cycle, an averaged variational principle whose Lagrangian density function depends on θ only through the first derivatives of θ (rather than on θ itself or any higher derivatives of θ); and thirdly, that the boundary conditions with respect to the cross-space variables are such that the cross-space integral of the phase-averaged Lagrangian density function reduce to zero when *actual solutions* of the homogeneous boundary-value problem (as opposed to the 'test functions' of variational calculus) are substituted into it.

This algorithm for deriving the conservation law for $\mathscr{L}_\omega$ is, in its essentials, the one followed above to derive (5.4), and the same process may be applied to derive Nayfeh's conservation law from a variational principle whose Euler equations are (6.4) and (6.5) above. An example of such a variational principle is

$$\delta \iiint_D \int_{t_0}^{t_1} \left(p^* \frac{\partial u_j}{\partial x_j} - p\frac{\partial u_j^*}{\partial x_j} + \tfrac{1}{2}(u_j^* \operatorname{lsd} u_j - u_j \operatorname{lsd} u_j^*) + u_2\left(u_1^* \frac{\partial U_1}{\partial x_2} + u_3^* \frac{\partial U_3}{\partial x_2}\right)\right) \mathrm{d}t\,\mathrm{d}V = 0, \quad (6.10)$$

in which the domain D has the same meaning as described in §4.1 above and all eight of the quantities u_j, p, u_j^*, p^*, $j \in \{1, 2, 3\}$ are independently variable. One must also assume that the quantities u_j and u_j^* are given prescribed values on the cylindrical surface $x_1^2 + x_3^2 =$ constant, that makes up the horizontal extremities of D and that

u_j and u_j^* are given prescribed values on the surfaces $x_2 = 0$ and $x_2 = h$. The small-disturbance equations (6.4*a*, *b*) follow by taking independent variations of each of the starred quantities, and the adjoint equations (6.5) follow by taking independent variations of the non-starred ones. If trial functions of the form

$$u_i = \tilde{u}_i \mathrm{e}^{\mathrm{i}\theta}, \quad p = \tilde{p}\,\mathrm{e}^{\mathrm{i}\theta}, \quad u_i^* = \tilde{u}_i^* \,\mathrm{e}^{-\mathrm{i}\theta}, \quad p^* = \tilde{p}^* \,\mathrm{e}^{-\mathrm{i}\theta}, \qquad (6.11\,a\text{–}d)$$

are substituted into (6.10), the result is

$$\delta \iiint_D \int_{t_0}^{t_1} \tilde{L}\,\mathrm{d}t\,\mathrm{d}V = 0, \qquad (6.12)$$

where

$$\tilde{L} = \tilde{p}^* \left(i\frac{\partial \theta}{\partial x_1}\tilde{u}_1 + \frac{\partial \tilde{u}_2}{\partial x_2} + i\frac{\partial \theta}{\partial x_3}\tilde{u}_3 \right) + \tilde{p}\left(i\frac{\partial \theta}{\partial x_1}\tilde{u}_1^* - \frac{\partial \tilde{u}_2^*}{\partial x_2} + i\frac{\partial \theta}{\partial x_3}\tilde{u}_3^* \right)$$

$$+ i\tilde{u}_j^* \tilde{u}_j \operatorname{lsd}\theta + \tilde{u}_2 \left(\tilde{u}_1^* \frac{\partial U_1}{\partial x_2} + \tilde{u}_3^* \frac{\partial U_3}{\partial x_2} \right). \qquad (6.13)$$

The Euler equation corresponding to independent variations with respect to θ is similar to (4.25) above, the only distinction being that (6.13) is the definition of $\tilde{L}$ rather than (4.16). Letting

$$-\frac{\partial \theta}{\partial t} = \omega, \quad \frac{\partial \theta}{\partial x_1} = k_1, \quad \frac{\partial \theta}{\partial x_3} = k_3,$$

and noting that

$$\operatorname{lsd}\theta \equiv \frac{\partial \theta}{\partial t} + U_1\frac{\partial \theta}{\partial x_1} + U_3\frac{\partial \theta}{\partial x_3} = -\omega + U_1 k_1 + U_3 k_3,$$

one obtains

$$\mathscr{L}_\omega \equiv \int_0^h \tilde{L}_\omega \,\mathrm{d}x_2 = -\mathrm{i}\int_0^h \tilde{u}_j \tilde{u}_j^* \,\mathrm{d}x_2.$$

Apart from the proportionality factor $-\mathrm{i}$, this expression agrees with Nayfeh's quantity (6.8) in the limit as h tends to infinity. The derivation of the conservation law for this $\mathscr{L}_\omega$ proceeds exactly as described above, thus confirming the consistency of Nayfeh's result with the variational formalism employed here.

Some discussion of the significance of the phrase 'wave action density' is in order. 'Action', in this context, refers to the integral in Hamilton's principle of mechanics, which for a single particle, is a time integral of a Lagrangian function $L = T - V$, T and V being kinetic energy and potential energy of the particle. In continuum mechanics, the time integral of a Lagrangian is replaced by a time–space integral of a Lagrangian density. There is great beauty and symmetry in the idea that a quantity that is 'stationary' in the mechanics of particles and continua should transform to a quantity that is 'conserved' in the case of wave packets and slowly varying wavetrains. The former quantity is related to the *time integral* of a Lagrangian density; the latter quantity is a *frequency derivative* of a (phase-averaged) Lagrangian density. The set of variables that we have employed in the present work was selected, in part, to exhibit this duality between ordinary mechanics and wave mechanics.

Apart from the considerations already mentioned, variational statements of the basic physics of a problem are often quite useful. The derivation of the differential equations of motion of a problem in 'generalized coordinates' is often accomplished more efficiently and reliably from a variational statement than by direct transformation of the differential equations in Cartesian coordinates. Some efficient numerical methods, such as the finite-element method, are predicated on the availability of a variational statement of the basic physics. It is possible, therefore, that simple variational statements of the physics of shear waves, such as those present herein, may lead to progress in numerical simulation techniques.

The author is indebted to his collaborator, C. E. M. Hailey, and his former mentor, M. T. Landahl, for many useful discussions on the subject matter of this paper.

REFERENCES

CHIN, W. C. 1980 *AIAA J.* **18**, 149.

DRAZIN, P. G. & HOWARD, L. N. 1966 *Adv. Appl. Mech.* **9**, 1.

ECKART, C. 1963 *Phys. Fluids* **6**, 1037.

GOLDSTEIN, H. 1980 *Classical Mechanics.* Addison-Wesley.

HAYES, W. D. 1970*a* *Proc. R. Soc. Lond.* A **320**, 187.

HAYES, W. D. 1970*b* *Proc. R. Soc. Lond.* A **320**, 209.

HILLE, E. 1969 *Lectures on Ordinary Differential Equations.* Addison-Wesley.

ITOH, N. 1980 In *Proc. IUTAM Symp. on Laminar–Turbulent Transition* (ed. R. Eppler & H. Fasel), pp. 86–95. Springer.

ITOH, N. 1981 *Proc. R. Soc. Lond.* A **375**, 565.

JIMENEZ, J. & WHITHAM, G. B. 1976 *Proc. R. Soc. Lond.* A **349**, 277.

KLEBANOFF, P. S., TIDSTROM, K. D. & SARGENT, L. M. 1962 *J. Fluid Mech.* **12**, 1.

LANDAHL, M. T. 1972 *J. Fluid Mech.* **56**, 775.

LANDAHL, M. T. 1982 *Phys. Fluids* **25**, 1512.

LIGHTHILL, M. J. 1978 *Waves in Fluids.* Cambridge University Press.

NAYFEH, A. H. 1980 In *Proc. IUTAM Symp. on Laminar–Turbulent Transition* (ed. R. Eppler & H. Fasel), pp. 201–217. Springer.

SELIGER, R. L. & WHITHAM, G. B. 1968 *Proc. R. Soc. Lond.* A **305**, 1.

WHITHAM, G. B. 1974 *Linear and Nonlinear Waves.* Wiley-Interscience.