Weakly nonlinear stability theory of stratified shear flows

By S. A. MASLOWE

Department of Mathematics, McGill University, Montreal, Quebec

(Received 6 September 1976; revised 11 March 1977. Communicated by Dr P. G. Drazin)

SUMMARY

A derivation is given of the first-order, nonlinear amplitude equation governing the temporal evolution of finite-amplitude waves in stratified shear flows. The theory has been developed on an essentially inviscid basis by perturbing away from the linear neutral stability curve in Richardson number–wavenumber space. However, viscosity and heat conduction are still required in order to eliminate the singularities that occur in the inviscid limit. Holmboe's mixing layer model has been studied by applying the present theory and the results show, surprisingly, that subcritical instability can occur, i.e. modes that would be stable on a linear basis (e.g. when the Richardson number is greater than \( \frac{1}{3} \)) become unstable when the initial perturbation amplitude is greater than some critical value. An instability due to resonance which occurs at a Richardson number of 0.22 is also revealed by the analysis. These results have interesting implications in connection with clear air turbulence which are discussed herein.

1. INTRODUCTION

Geophysical fluid dynamicists have recognized for some time now that linearized stability theory, while perhaps necessary, is not sufficient to describe adequately many of the more important wave-like phenomena occurring in the atmosphere and oceans. Non-linearity often plays an essential role in these phenomena. This is particularly true of the large-amplitude ‘Kelvin–Helmholtz billows’ that have been observed in the troposphere by sensitive radars and reported recently in the literature (see, e.g., Browning 1971; Reed and Hardy 1972; Hardy, Reed and Mather 1973). The amplitude of these billows can be several hundred metres and may, on occasion, be as large as one or two kilometres. Measurement of their phase speeds generally indicates the presence of a critical level, i.e. a height where the mean wind and wave speeds are equal. This is essential in order for energy to be transferred from the mean wind shear to the wave.

One would like to relate these observations to theoretical results when that is possible. This is done most easily if the theory employed is linear stability theory. However, a linear theory is simply not capable of describing some of the most important features of billow development. These include the detailed flow structure (which typically contains thin diffusive layers that may be locally unstable), the determination of the peak amplitude as well as any amplitude-dependent properties, and the modification of the mean flow as it transfers energy to the wave.

Probably the most interesting and controversial question in any discussion of the capabilities of linear theory concerns the ‘critical Richardson number’ (the local Richardson number is here denoted \( Ri \)). According to the well-known Miles–Howard criterion, a necessary condition for instability is that \( Ri \) be less than \( \frac{1}{4} \) at some point in the flow. Laboratory observations seem to support that conclusion; however, experiments to date have been conducted at Reynolds numbers several orders of magnitude smaller than those characteristic of the atmosphere and, therefore, cannot be regarded as definitive. Atmospheric measurements, on the other hand, often yield much larger values of \( Ri \) than one would expect on the basis of the linear criterion, the range 0.15 < \( Ri_{\text{min}} \) < 0.80 being typical.

This apparent discrepancy between theory and observation has often been attributed to inexactitude in the measurements. For example, Browning (p. 295) in discussing one case
for which the minimum value of $Ri$ was about 0.7 concluded that small-scale wind changes were responsible for 'the anomalously large value of $Ri$'. Similarly, Hooke, Hall and Gossard (1973, p. 36) in reporting their interesting acoustic sounder experiments, suggest that the observation of an unstable wave outside of the linear stability boundary was due to uncertainty in the data.

It is not our purpose here to argue that in these two cases the measurements were really correct, but rather to make the point that these results are not anomalous if one is willing to recognize the limitations of a linear theory. Indeed, the theoretical results presented below show quite clearly that a single finite-amplitude mode can be unstable when $Ri$ is greater than $\frac{1}{4}$. Moreover, the analysis reveals that two initially infinitesimal waves that are linearly stable can interact resonantly in such a way that both will amplify by extracting energy from the mean flow. Although only the neighbourhood of the linear stability boundary is investigated in this paper, there is no reason to doubt that such nonlinear instabilities are also possible at larger values of $Ri$.

The technique that has been used to obtain these results is the weakly nonlinear stability theory. What the latter term implies is a straightforward perturbation expansion in powers of an amplitude parameter $\varepsilon$; the expected result is a nonlinear amplitude equation from which one recovers the known result of linear theory when $\varepsilon = 0$. In spite of its limitation to small amplitudes, a weakly nonlinear theory does remove some of the more severe limitations of a linear analysis; e.g. unsteady perturbations are no longer constrained to grow or decay exponentially (for ever) and the initial amplitude is no longer merely an arbitrary constant. Such a theory has existed for homogeneous shear flows since 1960 when Stuart and Watson showed that the perturbation amplitude $A(t)$ satisfies an equation of the form

$$\frac{1}{A}(dA/dt) = a_0 + \varepsilon^2 a_2 |A|^2 + O(\varepsilon^3)$$

(1)

In a frame of reference moving with the wave speed $c$, the quantity $a_0$ can be identified with $\alpha c_1$, the amplification factor of linearized theory; however, it is the Landau constant $a_2$ that is of central interest in the nonlinear theory. Supposing, for the sake of simplicity, that $a_2$ is real, the following possibilities arise: (i) $a_2 < 0$ means that a linearly unstable perturbation ($a_0 > 0$) will evolve toward a steady finite-amplitude state having an equilibrium amplitude $|A_0|^2 = -(a_0/a_2 \varepsilon^2)$. This is termed the supercritical case; (ii) with $a_2 > 0$, modes that would be damped ($a_0 < 0$) in the linearized theory can now amplify if their initial amplitude satisfies the condition $|A(0)|^2 > -(a_0/a_2 \varepsilon^2)$. Such destabilization by finite perturbations is known as subcritical instability.

For a stratified shear flow in which the minimum value of the local Richardson number is, say, 0.15-0.35, i.e. not far from its critical value, $|a_0| \ll 1$ so that a weakly nonlinear theory ought to yield physically significant results throughout this important parameter range. Yet, the only examples of this type of investigation are the parallel studies of Drazin (1970) and Maslowe and Kelly (1970) on the discontinuous Kelvin–Helmholtz flow. By considering such discontinuous profiles, one avoids having to contend with the singular behaviour that is associated with the inviscid neutral modes that represent stability boundaries for continuous stratified shear flows. Both viscosity and heat conduction are required in order to remedy this singular behaviour which is too strong to be dealt with by analytic continuation into the complex plane as is done in the case of homogeneous shear flows.

Clearly, there are great computational advantages in studying broken-line profiles. The disadvantage is that one does not always know what significance can be attached to the results (or indeed if they have any significance in the nonlinear case). Whereas Drazin and Howard (1961) have shown that in the linear theory the results for the Kelvin–Helmholtz model coincide with the long-wave limit for any continuous flow whose velocity and density
profiles approach constant values as $y \to \pm \infty$, there is no corresponding result in the nonlinear theory. In fact, Drazin obtained an amplitude equation for the Kelvin–Helmholtz flow that is second-order in time, whereas the equation obtained herein for continuous flows is the first-order equation (1).

The approach to be employed can be best illustrated with a concrete example. We will therefore consider the mixing layer model first studied by Holmboe which consists of the nondimensional velocity profile $\bar{u} = \tanh y$ and the density profile $\bar{\rho} = \exp(-\beta \tanh y)$. The neutral eigensolution for this flow in the inviscid limit is given by $c = 0$ and $J_0 = \alpha(1 - \alpha)$, where $c$ is the perturbation phase speed, $\alpha$ the wavenumber and $J_0$ is an overall Richardson number.

From the numerical computations of Maslowe and Thompson (1971), it can be seen that the neutral curve for a Reynolds number $Re \geq 100$ is very close to the inviscid result. Hence, the inviscid stability boundary can be employed to illustrate the basic idea. As indicated in Fig. 1, we wish to perturb away from some point on the neutral curve such as A. Whether it will be more interesting to perturb in the direction of B or B' depends upon whether the Landau constant turns out to be negative or positive, respectively. (The peak of the neutral curve near $J_0 = \frac{1}{4}$ is the obvious point about which to perturb. However, the parameter $a_2$ is defined all along the neutral curve and it turns out to be very profitable to determine its complete behaviour.)

Prior to making the calculations described in this paper, it was expected that $a_2$ would be negative so that the supercritical case, i.e. B in Fig. 1, would be the point of interest. The reason for this expectation is that supercritical equilibrium states were obtained previously in the limiting cases $\alpha = 1, J_0 = 0$ (by Schade (1964); point C in Fig. 1) and in the Kelvin–Helmholtz limit $J_0 = \alpha \to 0$ (point D in Fig. 1). However, the latter solutions correspond to limit-cycle oscillations and cannot be compared directly with solutions of Eq. (1).

Contrary to expectations, $a_2$ turns out to be positive in the case of Holmboe's model; hence, points such as B' in Fig. 1 are susceptible to subcritical instability. In a previous study (Maslowe 1973), it was shown that when $J_0 > \frac{1}{4}$ the Taylor–Goldstein equation of linear inviscid theory admits a continuous spectrum of singular neutral modes provided that there is no phase change across the critical layer. The latter condition is compatible with an inviscid nonlinear critical layer. The present result $a_2 > 0$ lends considerable support to the conjecture that such modes do occur in the high Reynolds number shear flows produced in the atmosphere and oceans and that they are the result of a nonlinear amplification mechanism. Further discussion will be postponed until section 5 so that the theory can first be presented.

2. General theory

(a) Preliminary considerations

The point of view taken in this study is that the dynamics of the nonlinear modes
considered are governed primarily by inviscid processes; hence, the role of viscosity and heat conduction is simply to remedy the singular behaviour of inviscid neutral modes. The Reynolds number is therefore to be regarded as fixed at some large, but finite, value.

The dimensionless governing equations that are obtained after making the Boussinesq approximation are the vorticity equation

$$\frac{D}{Dt} \nabla^2 \psi - \frac{g \rho}{\rho \partial x} = \frac{1}{Re} \nabla^2 (\nabla^2 \psi)$$

(2)

the energy equation

$$\frac{DT}{Dt} = \frac{1}{Re Pr} \nabla^2 T$$

(3)

and the equation of state

$$\rho = 1 - \beta T_0 (T - 1).$$

(4)

Here, \( \psi \) is the stream function which is related to the velocity components \((u, v)\) by \( u = \psi_y \) and \( v = -\psi_x \). Quantities not previously defined are \( Pr \), the Prandtl number, \( T_0 \), a reference temperature, and \( \beta \), the coefficient of thermal expansion.

A perturbation approach is to be employed in which the stream function, temperature and density will be expressed as power series in \( \epsilon \) and expanded about their mean values; \( \epsilon^2 \) will be identified subsequently with the perturbation kinetic energy. After substituting the power series into Eqs. (2)–(4), a sequence of linear problems results which are to be solved in succession. The \( O(\epsilon) \) perturbation to the mean flow satisfies the usual equations of linearized theory; however, the stream function, for example, at this order takes the form

$$\psi^{(1)} = A(\tau) \Phi_1(y) e^{ix(x-c\tau)} + A^*(\tau) \Phi_1^*(y) e^{-ix(x-c\tau)},$$

where \( \tau = \alpha \epsilon^2 t \) and the real quantity \( c \) is the phase speed. Hence, the amplitude evolves on the slow time scale \( \tau \) so that the ‘two-timing’ method can be used in which the time derivatives are transformed according to

$$\partial / \partial t \rightarrow \partial / \partial \tau + \alpha \epsilon^2 \partial / \partial \tau$$

(5)

The choice \( \tau = \alpha \epsilon^2 t \) is dictated by the form of the Landau equation (1) and also by the manner in which \( \alpha c_1 \) varies with distance from the neutral curve. (Here, \( \alpha c_1 \) denotes the amplification factor that linear theory would predict at the same values of \( \alpha, J_0 \) and \( Re \) under consideration in the nonlinear problem, where there is no such quantity as \( c_i \).) It will be shown that the amplitude equation is given by

$$\epsilon^2 (1/4A)(dA/d\tau) = a_0 + \epsilon^2 a_2 A^2$$

(6)

Recalling that \( a_0 \) is proportional to \( \alpha c_1 \), it can be seen that the desired balance occurs when \( \alpha c_1 \) is \( O(\epsilon^2) \) (provided that \( a_2 \) turns out to be \( O(1) \), i.e. the theory is restricted to those values of \( J_0 \) and \( \alpha \) where that turns out to be the case).

The expansions for \( J_0 \) and \( \alpha \) can be determined in advance by employing Howard’s formula for perturbing away from the neutral curve. The relevant equations for Holmboe’s model, Eqs. (17) and (19) of Howard (1963), show that near \( \alpha = 0 \), for example, \( \alpha c_1 \) is proportional to \( J_{on} - J_0 \), where \( J_{on} \) is the Richardson number on the neutral curve. At the point \( \alpha = 1, J_0 = 0 \), on the other hand, Lin’s perturbation formula, or Eq. (19) of Howard, shows that \( \alpha c_1 \) is proportional to \( \alpha_n - \alpha \). In general, we can therefore write

$$J_0 \sim J_{on} + \epsilon^2 J_2 + \ldots, \quad \text{and} \quad \alpha \sim \alpha_n + \epsilon^2 \alpha_2 + \ldots$$

(7)

where \( J_{on} = J_0(A) \) and \( \alpha_n = \alpha(A) \) at a point such as A in Fig. 1. With the use of these expansions, it will be seen that the amplitude equation (6) arises from an orthogonality condition that is imposed as a result of the nonlinear forcing due to the interaction of the fundamental disturbance mode and its first harmonic.
We consider a parallel shear flow with mean velocity profile \( \bar{u}(y) \) and density profile \( \bar{\rho} = \exp\{-\beta r(y)\} \). It is convenient to introduce a coordinate system moving with the wave speed, so we set
\[
\psi = \int_{-\infty}^y \{\bar{u}(y) - c\} \, dy + e\psi(\theta, y, \tau), \quad \text{and} \quad \rho = \rho(y) + \epsilon \beta(\theta, y, \tau),
\]
where \( \theta = \alpha x \) and the temperature can be eliminated by using Eq. (4). On the fast time scale, the flow is steady in the moving system so that \( \partial / \partial t = 0 \). Substituting now into Eqs. (2)–(4), we find that the perturbation stream function and density satisfy
\[
e^2 \nabla^2 \psi + (\bar{u} - c) \nabla^2 \psi + \bar{u}'' \psi + j_0(\bar{r} / \beta) \bar{\rho} = 1/\alpha Re \nabla^2 (\nabla^2 \psi) \quad (8)
\]
and
\[
e^2 \bar{\rho} + (\bar{u} - c) \bar{\rho} - \bar{p} \psi + e(\psi \bar{\rho} - \psi \bar{\rho}_y) = 1/\alpha Re Pr \nabla^2 \bar{\rho}, \quad (9)
\]
where \( \nabla^2 = \alpha^2 \partial^2 / \partial x^2 + \partial^2 / \partial y^2 \) and \( j_0 = g \beta \). To solve Eqs. (8) and (9) the perturbation stream function is expanded as
\[
e^2 \psi \sim e^{\{\phi_1(\tau, y)e^{10} + \phi_2(\tau, y)e^{-10}\} + e^{2\{\phi_2(\tau, y)e^{210} + \phi_3e^{-210}\}} + e^{3\{\phi_3(\tau, y)e^{310} + \phi_3e^{310} + \phi_3e^{-310}\} + O(e^4)} (10)
\]
with a similar expansion for \( \bar{\rho} \).

The quantities \( j_0 \) and \( \alpha \) are also expanded according to Eq. (7). In the first-order problem, the variables are separated by writing \( \phi_1 = A(\tau) \Phi_1(\tau) \) and \( \rho_1(\tau, y) = A(\tau) P_1(\tau, y) \). It is possible to eliminate \( P_1 \) from the resulting two equations so that the problem for a linear neutral mode takes the form
\[
L_{(a)} \Phi_1 = (\bar{u} - c)^2 \nabla^2 \Phi_1 - \bar{u}''(\bar{u} - c) \Phi_1 + \frac{i}{\alpha Re} (\bar{u} - c)(1 + Pr^{-1}) \nabla^2 \Phi_1 +
\]
\[
+ j_0 \bar{r} \Phi_1 + \frac{1}{Pr (\alpha Re)^2} \nabla^2 \Phi_1 + \frac{i}{\alpha Re Pr} (2\bar{u}'' \nabla^2 \Phi_1 - 2\bar{u}''(\Phi_1 - \bar{u}^2) \Phi_1) = 0, \quad (11)
\]
where the operator \( \nabla^2 = d^2 / d y^2 - \alpha^2 \).

The solution of Eq. (11) with its associated homogeneous boundary conditions constitutes an eigenvalue problem wherein solutions only exist for suitable combinations of the parameters \( \alpha, \bar{r}, j_0, Re \) and \( Pr \). The numerical methods that have been employed to solve this problem are outlined in section 3.

Proceeding now to the \( O(e^2) \) problem, the variables can again be separated if we take \( \phi_2 = A^2 \Phi_2(\tau, y) \) and \( P_2 = A^2 P_2(\tau, y) \). Substituting into Eqs. (8) and (9) leads to the following pair of coupled equations:
\[
(\bar{u} - c) \nabla^2 \Phi_2 - \bar{u}'' \Phi_2 + j_0(\bar{r} / \beta) P_2 + (i/2 Re) \nabla^2 \Phi_2 = \frac{1}{4}(\Phi_1 \nabla^2 \Phi_1 - \Phi_1 \nabla^2 \Phi_1) \quad (12)
\]
and
\[
(i/2 Re Pr) \nabla^2 \Phi_2 = \frac{1}{4}(\Phi_1 \nabla^2 \Phi_1 - \Phi_1 \nabla^2 \Phi_1) \quad (13)
\]
where \( \nabla^2 \Phi_2 = d^2 / d y^2 - 4 \alpha^2 \). For numerical purposes, this is the most convenient form for these equations. However, it is significant, as will be seen later, that Eqs. (12) and (13) can be combined into the following single nonhomogeneous equation for \( \Phi_2 \):
\[
L_{(2a)} \Phi_2 = (\bar{u} - c)(\Phi_1 \nabla^2 \Phi_1 - \Phi_1 \nabla^2 \Phi_1) + (i/4 Re Pr) (\Phi_1 \nabla^2 \Phi_1 - \Phi_1 \nabla^2 \Phi_1) \quad (14)
\]
where the operator \( L_{(2a)} \) is defined by Eq. (11) with \( \alpha \) replaced everywhere by \( 2\alpha \).
There is also a mean flow distortion that occurs at $O(\varepsilon^2)$; however, this effect is probably not important for free shear layers at very high Reynolds numbers. Its neglect is consistent with the parallel flow approximation as discussed in section 5.

At $O(\varepsilon^3)$, the slow time variation first enters into the analysis and the amplitude equation will emerge as a result of the solvability condition to be imposed upon $\phi_{31}$. Writing $\phi_{31} = A^* A \Phi_{31}$, and eliminating the density, as before, leads to the equation

$$
A^* A L(\varepsilon) \Phi_{31} = \frac{dA}{d\varepsilon} (\bar{u} - c) \nabla^2 \Phi_{11} + \frac{J_{0n}}{\beta} P_1 + \frac{i}{\alpha Re Pr} \nabla^2 \Phi_{11} +
$$

$$+ (2(\bar{u} - c) x_2 y_2 - J_2 \bar{r}^2 + O(Re^{-1})) \Phi_{11} A - G(y)A^2 A^*, \quad (15)
$$

where

$$
G(y) = \left[ 2 \Phi_{11}^* \nabla^2 \Phi_{22} + \Phi_{11}^* \nabla^2 \Phi_{22} - (2 \Phi_{22}^* \nabla^2 \Phi_{11}^* + \Phi_{22}^* \nabla^2 \Phi_{11}^*) \right] (\bar{u} - c) +
$$

$$+ \frac{J_{0n}}{\beta} \left[ (P_2 \Phi_{11}^* - \Phi_{22}^* P_1) + (P_2 \Phi_{11}^* - \Phi_{22}^* P_1) \right] +
$$

$$+ \frac{i}{\alpha Re Pr} \left[ 5 \Phi_{11}^* \nabla^2 \Phi_{22} + 2 \Phi_{11}^* \nabla^2 \Phi_{22} + 2 \Phi_{22}^* \nabla^2 \Phi_{11}^* - 4 \Phi_{11}^* \nabla^2 \Phi_{22} - 4 \Phi_{11}^* \nabla^2 \Phi_{22} +
$$

$$- \Phi_{11}^* \nabla^2 \Phi_{22} + 3 \alpha^2 \left( \Phi_{11}^* \nabla^2 \Phi_{11}^* + 2 \Phi_{22}^* \nabla^2 \Phi_{11}^* \right) \right].
$$

The quantity $G(y)$ represents the nonlinear interaction of the second harmonic ($e^{2i\theta}$) and the fundamental mode ($e^{-i\theta}$) terms.

A necessary and sufficient condition for the existence of a solution to Eq. (15) is that the right side be orthogonal to the solution of the adjoint problem. The adjoint function $\chi$ satisfies, in the present case, the differential equation (cf. Eq. (11))

$$
(\bar{u} - c)^2 \nabla^2 \chi + 4 \bar{u}'(\bar{u} - c) \chi' + \frac{2i}{\alpha Re} (2 + P r^{-1}) \bar{u} \nabla^2 \chi' +
$$

$$+ \frac{i}{\alpha Re} \left[ (1 + P r^{-1})(\bar{u} - c) \nabla^2 \chi + 2 \bar{u}' \nabla^2 \chi + 4(\bar{u}' \chi') + \bar{u} \chi \right] -
$$

$$- \frac{1}{Pr(\alpha Re)^2} \nabla^2 \chi + \left[ \bar{u}''(\bar{u} - c) + 2 \bar{u}'^2 + J_0 \bar{r} \right] \chi = 0. \quad (16)
$$

The homogeneous boundary conditions satisfied by $\chi$ are given in section 3. Imposing now the orthogonality condition, we obtain

$$
\frac{dA}{d\varepsilon} \int_{-\infty}^{\infty} \{ (\bar{u} - c) \nabla^2 \Phi_{11} + \frac{J_{0n}}{\beta} P_1 + \frac{i}{\alpha Re Pr} \nabla^2 \Phi_{11} \} \chi dy +
$$

$$+ A \int_{-\infty}^{\infty} \{ 2(\bar{u} - c) x_2 y_2 - J_2 \bar{r} \} \Phi_{11} \chi dy - A^2 A^* \int_{-\infty}^{\infty} G(y) \chi dy = 0, \quad (17)
$$

which is simply the Landau equation (1). The numerical solution of Eqs. (11)–(17) will determine the coefficients $a_0$ and $a_2$ that appear in Eq. (1).

3. Numerical techniques

The eigenvalue problem associated with Eq. (11) has been solved using a 4th-order Runge–Kutta subroutine and double-precision complex arithmetic. By exploiting the symmetry of the mean flow profiles in Holmboe's case it is possible to reduce the range of
integration to the semi-infinite domain $-\infty \leq y \leq 0$. Taking the real part of $\Phi_1$ to be even and the imaginary part odd leads to the following boundary conditions at $y = 0$:

$$
\Phi_1'(0) = \Phi_1''(0) = 0 = \Phi_1(0) = \Phi_1'(0) = \Phi_1''(0) = 0.
$$

Also, from the symmetry of the equations one concludes that $c_r = 0$.

The numerical integration is initiated at $y = -3$ by employing the asymptotic form of $\Phi_1$:

$$
\Phi_1 = B_1 \exp(\alpha y) + B_2 \exp(\alpha(1 + i \text{Re} \alpha) y) + B_3 \exp(\alpha(1 + i \text{Re} \alpha) y)
$$

as $y \to -\infty$

which is obtained by setting $\bar{u} = -1$ and $\bar{w} = \bar{z} = 0$ in Eq. (11). Three linearly independent solutions are obtained by successively setting two of the three constants $B_1$, $B_2$, and $B_3$ equal to zero while the remaining constant is set to one. These solutions can then be superimposed with the constants $B_1$ chosen such that all but one of the boundary conditions are satisfied; the remaining condition will be satisfied only when $\alpha$, $c_r$, $J_o$, $\text{Re}$ and $\text{Pr}$ are eigenvalues. This method is not suitable for very high Reynolds numbers, but was found to be adequate for $\alpha \text{Re} \text{Pr} \leq 150$. For further details the reader is referred to the paper of Maslowe and Thompson.

It should be noted at this point that the magnitude and variation with $\alpha$ of the Landau constant is strongly influenced by the normalization used for the arbitrary constant $B_1$. The procedure employed here is based upon a consideration of the perturbation kinetic energy in the fundamental mode. This is given by

$$
E(t) = e^2 A^2 \int_{-\infty}^{\infty} (|\Phi_1|^2 + \alpha^2 |\Phi_2|^2) dy.
$$

All results are referred to the ‘standard solution’ $\alpha = 1$, $\Phi_1 = \text{sech} y$, corresponding to the case $J_0 = 0$ and $\text{Re} = \infty$. For that case the integral in Eq. (18) is equal to $8/3$; the constant $B_1$, was accordingly adjusted so that the value of this integral was always $8/3$. If, in addition, we always set $A(0) = \frac{1}{2} \sqrt{6}$, then $e^2$ is equal to the initial kinetic energy of the linear perturbation.

Returning now to the topic of numerical procedures, the solution of Eq. (14) turns out to be rather difficult in comparison to Eq. (11). The difficulties are due to the following features of the problem: (i) the Reynolds number is effectively doubled because of the appearance of $2 \alpha \text{Re}$ instead of $\alpha \text{Re}$ in the operator $L_{(2\alpha)}$; and (ii) the nonhomogeneity of the equation coupled with the unboundedness of the flow makes it difficult to separate out the homogeneous and particular solutions. As a result, the asymptotic behaviour for $y \to -\infty$ cannot be readily obtained.

The whole matter can be resolved very nicely through the use of a finite-difference method (see e.g. section 17.8 of the text by Carrier and Pearson 1968). The latter method is not susceptible to the ‘contamination’ sort of instability occurring in the solution of ODEs that have a small parameter multiplying the highest-ordered derivative and has the additional advantage that the boundary conditions are easily imposed.

In the procedure employed, Eqs. (12) and (13) were rewritten as a coupled set of six real second-order ODEs. The second derivatives were then expressed in difference form according to

$$
\Phi''(y_n) \simeq (\Phi_{n+1} - 2\Phi_n + \Phi_{n-1})/h^2 + O(h^2),
$$

where the step size $h$ was taken to be 0.05. The boundary conditions imposed at $y = 0$ are

$$
\Phi_2'(0) = \Phi_2''(0) = \Phi_2(0) = \Phi_2'(0) = P_2(0) = P_2'(0) = 0.
$$

At $y = -3$, it was found that homogeneous boundary conditions could be used without any
significant loss of accuracy due to the rapid exponential decay \( \exp(2\alpha y) \) of \( \Phi_2 \) and \( P_2 \) as \( y \to -\infty \). A more detailed discussion of these matters can be found in the paper by Maslowe (1977) dealing with the unstratified counterpart of the present problem.

The numerical solution for \( \Phi_2 \) is of interest because of the strong critical-point singularity that occurs at \( y = 0 \) in the inviscid theory. Whereas \( \Phi_1 \sim y^{1-\alpha} \) as \( y \to 0 \) for an inviscid neutral mode, \( \Phi_2 \sim y^{-2\alpha} \) as \( y \to 0 \). Of course, including viscosity and heat conduction removes the singularity, but at high values of \( Re \), \( |\Phi_2| \gg |\Phi_1| \) so that the theory becomes limited to quite small values of \( \epsilon \). This effect is clearly illustrated by the results in Fig. 2(a) for the case \( \alpha = 0.5 \), \( Re = 100 \); for larger values of \( Re \), the breakdown is more dramatic. By contrast, \( |\Phi_2| < |\Phi_1| \) in the case of homogeneous flow (see Fig. 2(b)), where the inviscid weakly nonlinear theory breaks down at \( O(\epsilon^3) \) instead of \( O(\epsilon^2) \).

Similarly, the adjoint function \( \chi \) is singular in the inviscid theory behaving as \( y^{-\alpha} \) for \( y \ll 1 \). A solution of Eq. (16) is shown in Fig. 3, where the boundary conditions for \( \chi \) are the same as those for \( \Phi_1 \), i.e. \( \chi \) was taken to be even and \( \chi_1 \) odd.
4 Computed Results and Some Generalizations of the Theory

(a) Coefficients of the amplitude equation

The integrals appearing in Eq. (17) were evaluated using Simpson's rule after first computing and storing the functions $\Phi_1$, $\Phi_2$, $P_1$, $P_2$ and $\chi$. To assist in interpreting the results let us define the integrals

\[
I_1 = -i \int_{-\infty}^{\infty} \left[ (\bar{u} - c) \nabla_x^2 \Phi_1 + \frac{J_{0n}}{\beta} P_1 + \frac{i}{\alpha Re Pr} \nabla_x^2 \Phi_1 \right] \chi \, dy,
\]

\[
I_{21} = 2 \int_{-\infty}^{\infty} (\bar{u} - c)^2 \Phi_1 \chi \, dy, \quad I_{22} = -\int_{-\infty}^{\infty} \bar{v}' \Phi_1 \chi \, dy \quad \text{and} \quad I_3 = \int_{-\infty}^{\infty} G \chi \, dy.
\]

By taking into account the symmetry of the real and imaginary parts of their integrands, we conclude that $I_1, I_{21}, I_{22}$ and $I_3$ are all real numbers. The nonzero part of the above integrals, being even, can be evaluated by integrating numerically from $y = -\infty$ to $y = 0$ and doubling the result.

These integrals can be related to more familiar quantities by rewriting Eq. (17) in terms of the original fast time scale and substituting from Eq. (7) for $J_2$ and $\alpha_2$ to obtain

\[
\frac{1}{A} \frac{dA}{dt} = \alpha_n^2 \frac{I_{21}}{I_1} (\alpha - \alpha_n) + \alpha_n \frac{I_{22}}{I_1} (J_0 - J_{0n}) - \alpha_n e^{2I_3/I_1} |A|^2. \tag{19}
\]

The first two terms represent the value of $\alpha c_i$ in the linear theory that would be obtained by expanding $\alpha c_i$ in a Taylor series about a point on the neutral stability curve, i.e.,

\[
\alpha_n^2(I_{21}/I_1) = [\partial(\alpha c_i)/\partial \alpha]_{J_0, Re} \quad \text{and} \quad \alpha_n(I_{22}/I_1) = [\partial(\alpha c_i)/\partial J_0]_{\alpha, Re}.
\]

When $Re = \infty$, the above quantities can be expressed in terms of beta functions according to Howard's perturbation formula. At $Re = 100$, their magnitudes are reduced by about 15% compared with the inviscid values.

Of primary interest in the present study is the Landau constant which is given by

\[
a_2 = -\alpha_n(I_3/I_1). \tag{20}
\]

this result following directly from a comparison of Eq. (19) with Eq. (1). In Fig. 4, the Landau constant is shown as a function of $\alpha$ with $Re = 100$ and $Pr = 0.72$. For a homogeneous tanh $y$ shear layer, $a_2 = -1.504$ at $Re = 100$ as indicated in the figure; surprisingly,
a small amount of stratification causes $a_2$ to become positive and, as $\alpha$ is made smaller, the magnitude of $a_2$ becomes substantial. Qualitatively similar results are illustrated in Fig. 5 where $\alpha Re$, instead of $Re$, is held constant at a value of 100.

![Figure 5. Variation of the Landau constant with $\alpha$ while moving along the neutral curve at constant $Re = 100$.](image)

From a physical point of view, it is more interesting to see these results presented as in Fig. 6 where for two Richardson numbers, 0.25 and 0.30, the critical amplitude required for destabilization is shown as a function of $\alpha$. Recalling our convention that $A(0)^2 = \frac{3}{8}$, the critical amplitude is computed from $\varepsilon^2 = -(8/3)\alpha \epsilon_i/a_2$, where $\alpha \epsilon_i$ was found directly from Eq. (11) rather than Eq. (19). It is seen that for $J_0 = 0.25$ the critical values of $\varepsilon$ are sufficiently small that the conditions for validity of the weakly nonlinear theory are easily satisfied, whereas at $J_0 = 0.30$ the situation appears to be marginal. As $J_0$ increases, the preferred wavenumber, corresponding to the minimum value of $\varepsilon_{crit}$, shifts toward higher values due to the variation of $\alpha \epsilon_i$ with $J_0$.

![Figure 6. Critical amplitude for finite-amplitude instability to occur under linearly unstable conditions.](image)

The reader will notice that in both Figs. 4 and 5 $a_2$ is singular for $\alpha \approx 0.31$. This breakdown of the theory is due to a resonance that in the inviscid limit occurs at $J_0 = 2/9$ and involves neutral modes having wavenumbers of $\frac{1}{4}$ and $\frac{1}{3}$. Kelly (1968) pointed out that such a resonance could occur in Holmboe's model, but did not apply his theory to that case because of the necessity to add diffusive effects in dealing with the singular neutral modes. The significance of Eq. (14) is now clear; because the r.h.s. of this equation is generally nonzero, and there is one value of $J_0$ for which both $x$ and $2x$ are eigenvalues of Eq. (11), solutions of Eq. (14) are not obtainable for that particular $J_0$ and $x$. The theory associated with this exceptional case is outlined below. First, however, one additional set of results for $a_2$ will be discussed.
Fig. 7. Variation of the Landau constant with Reynolds number at a fixed wavenumber of $\alpha = 0.4$.

Fig. 7 illustrates the variation of $a_2$ with $Re$ at a constant value of $\alpha = 0.4$. One expects that $a_2 \sim Re^\delta$ as $Re$ becomes large due to the following theoretical considerations. The inviscid linear theory, which is singular for neutral modes, follows from the limit $\varepsilon \to 0$ and $Re \to \infty$ in Eqs. (8) and (9). It has been implicit throughout this analysis that the appropriate remedy for this singular behaviour is to include diffusive effects. However, an alternative is to allow nonlinearity to dominate within a critical layer whose thickness is given by $\varepsilon_p$, where $\frac{1}{2} \leq p \leq \frac{3}{2}$ corresponding to $0 \leq J_0 \leq 1$. The latter approach was applied to Holmboe’s model by Maslowe (1973). The linear diffusive approach is only applicable when the parameter $\lambda = (\alpha Re)^{\frac{\delta}{p}} > 1$. If Eqs. (8) and (9) are rescaled using the viscous critical layer thickness, $(\alpha Re)^{-\frac{\delta}{p}}$, then the nonlinear terms are multiplied by $\lambda^{-\frac{\delta}{p}}$. Hence, the Landau constant, which also multiplies the nonlinear terms (in the amplitude equation), will behave as $\lambda^{-\frac{\delta}{p}}$, or equivalently $Re^{\frac{\delta}{p}}$, when the weakly nonlinear theory breaks down. The numerical results in Fig. 7 follow quite closely that behaviour.

(b) Resonant interaction of neutral modes

The linear solution for this case consists of the following superposition of two neutral modes:

$$\psi^{(11)} = A_1(\tau)\Phi_1 e^{i\theta} + A_1^* \Phi_1^* e^{-i\theta} + A_2(\tau)\Phi_2 e^{2i\theta} + A_2^* \Phi_2^* e^{-2i\theta}.$$

An important difference vis-à-vis the single-mode case is that now $\tau = \varepsilon \tilde{t}$ so that the instability occurs on a faster time scale. At the next order in $\varepsilon$, secular terms will arise due to the interactions

$$e^{i\theta} e^{i\theta} = e^{2i\theta} \text{ and } e^{2i\theta} e^{-i\theta} = e^{i\theta}.$$

It is found that the variables can be separated if $A_1$ and $A_2$ satisfy the equations

$$\frac{dA_1}{d\tau} = \gamma_1 A_1^* A_2 \quad \text{and} \quad \frac{dA_2}{d\tau} = \gamma_2 A_2^2,$$

where the constants $\gamma_1$ and $\gamma_2$ are determined by imposing orthogonality conditions on the $O(\varepsilon^2)$ terms. The detailed analysis and computational results will be presented at a future date; however, preliminary results show that both $\gamma_1$ and $\gamma_2$ are real negative numbers and that $|\gamma_1| \approx 30|\gamma_2|$. Consideration of the scaling used in the resonant case indicates that the single-mode analysis probably breaks down when $|\alpha - \alpha_c| \sim O(\varepsilon)$ or less.

Eqs. (21) are characteristic of second-order triad resonant interactions and their solution can be obtained in terms of elliptic integrals. The properties of the solution to Eqs. (21) have been discussed by McGoldrick (1970) in a study of capillary–gravity waves. However, there are important differences in the present case; unbounded growth
can now occur because of the energy contained in the mean flow, whereas the process in the case of capillary–gravity waves is one of energy sharing.

To indicate the various possibilities, we introduce the real quantities \( a \) and \( \phi \), the amplitude and phase, by writing \( A = \frac{1}{2}ae^{-i\phi} \) and \( A^* = \frac{1}{2}ae^{i\phi} \). In terms of \( a \) and \( \phi \), Eqs. (21) become

\[
\begin{align*}
\frac{da_1}{d\tau} &= \frac{1}{2}y_1a_1a_2\cos \theta, \quad (22a) \\
\frac{da_2}{d\tau} &= \frac{1}{2}y_2a_1^2\cos \theta, \quad (22b) \\
\frac{d\phi_1}{d\tau} &= \frac{1}{2}y_1a_1a_2\sin \theta, \quad (22c) \\
\frac{a_2}{a_1}\frac{d\phi_2}{d\tau} &= \frac{1}{2}y_2a_1^2\sin \theta, \quad (22d)
\end{align*}
\]

where \( \theta = 2\phi_1 - \phi_2 \) is termed the relative phase.

From Eqs. (22a) and (22b), one can derive the energy integral

\[
a_1^2 - (y_1/y_2)a_2^2 = E, \quad (23)
\]

where \( E \) is a constant. This important result shows that both modes can amplify at the same time by extracting energy from the mean flow (recall that \( y_1/y_2 \) is positive). An integral involving the relative phase can also be derived from Eqs. (22a)–(22d), namely

\[
a_1^2a_2\sin \theta = L. \quad (24)
\]

Finally, the relationships (23) and (24) involving the constants \( E \) and \( L \) can be used to derive the following equation for \( a_1^2 \):

\[
\frac{d^2a_1^2}{d\tau^2} - \frac{1}{2}y_1y_2a_1^2(3a_1^2 - E) = 0. \quad (25)
\]

A phase plane study of Eq. (25) shows that \( a_1^2 \) will become unbounded in time provided that \( \pi/2 < \theta < 3\pi/2 \). This is true whether \( E \) is positive or negative, because the solution trajectories all lie to the right of a saddle point which is located at the origin when \( E \leq 0 \) or at \( a_1^2 = E/3 \) if \( E > 0 \). The importance of the relative phase can be seen clearly in the numerical calculations presented in Fig. 16 of Patnaik, Sherman and Corcos (1976). Their results were obtained by solving the Boussinesq equations with a disturbance consisting of a long and short wave having the wavenumbers \( \alpha = 0.215 \) and \( \alpha = 0.43 \). Other parameters were \( J_0 = 0.07 \), \( Re = 50 \) and \( Pr = 0.72 \). Two cases were run, corresponding to different values of \( \theta \); the nature of the interaction was observed to be quite different in the two cases, although the long wave proved to be dominant each time. The latter observation is consistent with the result \( |y_1| \gg |y_2| \) obtained in the present study and with experimental observation.

(c) Wave trains

The foregoing results can easily be extended to describe the evolution of the envelope for a slowly varying wave train. From the discussion of Benney and Maslowe (1975), it is straightforward to infer that the amplitude evolution equation in the present case will take the form

\[
\begin{align*}
\mu \left[ \frac{\partial A}{\partial T} + \frac{\partial \omega}{\partial k} \frac{\partial A}{\partial X} \right] - \frac{1}{2} \mu \omega^2 \frac{\partial^2 A}{\partial X^2} &= \frac{\partial \omega}{\partial \lambda}(J_0 - J_{0n})A + \epsilon a_2 A^2 A^*.
\end{align*}
\]

where \( X = \mu x \) and \( T = \mu t \) are slow space and time variables. For Holmboe's flow, \( \partial \omega/\partial k \) is generally imaginary and \( \mu = \mu^2 \), so that the relevant amplitude equation is (26) with the \( O(\mu^{\alpha}) \) dispersion term neglected. However, at the peak of the neutral stability curve (approximately \( \alpha = 0.46 \) when \( Re = 100 \)) \( \partial \omega/\partial k = 0 \) and the appropriate scaling is \( \mu = \epsilon, X = \mu x \) and \( T' = \mu T = \mu^2 t \). Eq. (26) now becomes
\frac{\partial A}{\partial T} + \frac{1}{2} \frac{\partial^2 \omega}{\partial k^2} \frac{\partial^2 A}{\partial X^2} = - \frac{\partial \omega}{\partial J_0} J_2 A + a_2 A^2 A^*.

(27)

This equation is characteristic of many problems in hydrodynamic stability: when \( J_2 = 0 \), it resembles the nonlinear Schrödinger equation which arises frequently in nonlinear wave studies. Of primary interest in the present investigation is the case where all coefficients are real with \( a_2 > 0 \) and \( \partial^2 \omega / \partial k^2 < 0 \). For that case, numerical solutions of Eq. (27) were obtained by Hocking, Stewartson and Stuart (1972; section 7). They found that for certain initial conditions, \( A \) became unbounded locally due to nonlinear focusing. Although the coefficient of \( A_{xx} \) employed in their calculations corresponds closely with the present case, \( a_2 \) is much larger here than the value \( a_2 = 1 \) used by Hocking et al. It therefore seems likely that the tendency toward unbounded growth will be more pronounced for our stratified shear flow.

5. Discussion

The investigation described above has revealed several nonlinear mechanisms capable of producing subcritical instability in stratified shear flows. Although the mathematical requirements imposed are only satisfied strictly in the range 0.20 \( \leq J_0 \leq 0.30 \) the physical mechanisms themselves may well be operational outside that range. Instability, of course, does not necessarily lead to turbulence and the outcome might instead be some quasiperiodic structure of rather large amplitude.

Of the various instabilities explored in this paper, the resonant interaction mode would appear to be the one that is most observable. This belief is derived from the following considerations: (i) the amplification takes place on a faster time scale (\( \varepsilon^2 t \) as opposed to \( a^2 t \) for a single subcritical mode); (ii) interactions of this sort have been observed in experiments with homogeneous shear layers (e.g. those of Miksad (1972)); (iii) the mechanism does not require a finite perturbation to get started; and (iv) the resonance conditions do not have to be satisfied exactly for a strong interaction to take place. Examples of the latter situation are the already cited results of Patnaik et al. and the near-resonant example considered in section 4 of the paper by Kelly (1967).

Having speculated about the relative importance of the instability mechanisms, we now turn our attention to some of the limitations of the theory. Firstly, it should be recalled that the r.h.s. of Eq. (1) contains only the first two terms of a series that is slowly convergent in the case of greatest interest, i.e. when both terms are the same order of magnitude. In the case of plane Poiseuille flow, where \( a_2 \) is also greater than zero, Nishioka, Iida and Ichikawa (1975) have recently observed experimentally that subcritical instability is followed by turbulence. However, there is no reason to believe that such will always be the case; transition in free shear layers, in particular, is qualitatively very different from that in bounded flows. For stratified shear flows, it seems reasonable to conjecture that subcritical instability will lead to some new equilibrium state, possibly beyond the mathematical range of weakly nonlinear theory. The stability of that state would then dictate the eventual fate of this 'Kelvin–Helmholtz billow'. Should the amplitude become large enough so that the parameter \( \lambda = (\alpha Re \varepsilon^2)^{-1} \) is much less than one, nonlinear critical layer concepts would become relevant and localized instabilities might occur as suggested by Maslowe (1973). In any case, more terms in the Landau equation would be required to describe accurately the evolution of a subcritical mode.

Two effects that are related to each other and have been neglected in this analysis should be commented upon, namely, the distortion of the mean flow and the growth of the shear layer thickness. Although these effects are not dominant in the weakly nonlinear regime,
they do cause some reduction in the value of $a_2$ and could be significant if the theoretical results were being compared with laboratory experiments at moderate Reynolds numbers. However, it is highly unlikely that $a_2$ would change sign given its magnitude (cf. Figs. 4, 5 and 7). Moreover, $a_2$ increases with $Re$, while the 'nonparallelness' and distortion of the flow decrease, so it is clear that subcritical instability would still occur at a sufficiently high Reynolds number.

From the point of view of applicability to the atmosphere, the principal limitation of the weakly nonlinear approach is the condition $\lambda \gg 1$. Suppose we imagine an amplifying wave at fixed $Re$ whose initial amplitude is small enough so that the above condition is satisfied at $t = 0$. As the amplitude increases and $\lambda$ becomes of $O(1)$, the Landau 'constant' $a_2$ becomes a function of $\lambda$, so the theory breaks down. (In the terminology of stability theory, the phase change across the critical layer is reduced as $\lambda$ becomes smaller.) For very small values of $\lambda$, Eq. (1) may no longer even be the governing amplitude equation. It is quite possible, at least for modes near the stability boundary, that the amplitude equation will be second-order in time in the special case of an antisymmetric shear layer with $c = 0$. This is what occurs in the case of a homogeneous tanh $y$ shear layer (see Eq. (4.23) of Benney and Maslowswe). Clearly, there are many extensions to the theory that are needed before the evolution of a subcritical mode can be described fairly completely. The regime $\lambda \sim O(1)$ and smaller, with time dependence, seems especially pertinent (and difficult).

It would seem that most of these matters could be greatly clarified by numerical investigation of the full nonlinear equations. That is only true to some extent, however, because numerical schemes have yet to be developed that can deal adequately with high Reynolds number flows. Numerical instabilities at large times and resolution difficulties limit the scope of such investigations at present. A further difficulty is that in the atmosphere and oceans, waves are more likely to grow in space (or space and time) than in time. Whereas analytical methods can be employed with only a moderate increase in complexity to deal with spatially growing waves, numerical schemes seem to be limited by practical considerations to waves that are spatially periodic. Nonetheless, numerical simulations are able to provide some valuable information, the work of Patnaik et al. being a noteworthy example, and improved computational methods could greatly expand their future capability.

ACKNOWLEDGMENTS

The author is indebted to Professor R. E. Kelly of UCLA for his continued interest this work and for originally suggesting this problem a number of years ago. He has also benefited from many helpful discussions with Professor D. J. Benney of MIT and some perceptive observations by Dr M. A. Weissman of the National Physical Laboratory concerning the resonant case. Discussions with Dr K. R. Hardy concerning the radar observations were most helpful and his encouragement is greatly appreciated. Finally, he is pleased to acknowledge the capable programming assistance of Mr Richard Dubois of McGill University.

This research was supported by the US National Oceanic and Atmospheric Administration and the National Research Council of Canada.

REFERENCES


