The non-linear interaction of two disturbances in the thermal convection problem

By LEE A. SEGEL
Rensselaer Polytechnic Institute, Troy, N.Y.

(Received 13 March 1962)

In the thermal convection problem with free boundaries, the interaction of two 'roll' disturbances is considered. The problem is reduced to a pair of non-linear ordinary differential equations, which should also provide a model for the interaction of two disturbances in more general situations than that for which these equations have been derived. The equations contain several parameters which necessitates a discussion of various possible types of solution. Some representative results are: (1) under certain circumstances, an equilibrium state may be composed of a mixture of a linearly stable disturbance and a linearly unstable disturbance; (2) for the thermal convection problem, when the Rayleigh number is slightly above the minimum critical value, the equilibrium state will contain only one of two linearly unstable disturbances. These and other results are compared with experimental observations.

1. Introduction and principal conclusions

If a horizontal layer of fluid is slowly heated from below, when the Rayleigh number reaches a certain critical value it is a classical result of the linear theory of hydrodynamic stability (Pellew & Southwell 1940) that the original state of no motion and a linear temperature gradient becomes unstable and a cellular motion ensues. Linear theory incorrectly predicts that the velocities will increase exponentially with time and cannot predict which of an infinite number of possible cell shapes, all with the same amplification rate, will occur. Recent non-linear analyses, however, have provided a good understanding of the first point and at least a start on understanding the second. (See, for example, Malkus & Veronis 1958 and Segel & Stuart 1962.)

As soon as the critical Rayleigh number is exceeded, which it must be—even if slightly—in any physical situation, linear theory allows each of the previously mentioned cell shapes to occur in a continuous spectrum of sizes and corresponding amplification rates. Unlike the non-uniqueness *at* the critical Rayleigh number, the non-uniqueness of the type found when the critical Rayleigh number is exceeded occurs in all stability problems. For example, there is a spectrum of linearly unstable disturbances in flow between rotating cylinders when the critical Taylor number is exceeded and in boundary-layer flow when the critical Reynolds number is exceeded. Nevertheless, in the Rayleigh and Taylor cases the spectrum of the disturbance appears to contain a single sharp peak at a certain wavelength.

What apparently happens is that the non-linear terms somehow act to damp all but a very narrow band of the linearly unstable disturbances.

It is one of the principal objects of this paper to make a start towards explaining how this comes about by considering the non-linear interaction of two disturbances both of which are unstable by linear theory. (We also briefly consider interactions when one or both disturbances are linearly stable.) The central part of the analysis is a general discussion of the various types of behaviour possible for solutions to a pair of coupled non-linear ordinary differential equations (2.15*a*, *b*). These equations are a model for the interaction of two unstable disturbances in situations where the dominant effect is alteration of the disturbances' initial exponential growth in time, so that the results should apply to a variety of physical situations (see Stuart 1961). Let us emphasize, however, that in this paper the equations (2.15*a*, *b*) are rigorously (though formally) derived by applying a certain expansion procedure to the equations of the thermal instability problem. The disturbances are specified to be 'rolls' periodic in one horizontal direction and independent of the perpendicular horizontal direction, so that the cell-shape non-uniqueness problem is by-passed. Both horizontal boundary planes are taken to be free surfaces; this physically unrealistic assumption greatly simplifies the calculations. As is usual in thermal instability problems, the results are expected to have qualitative significance.

We first consider disturbances which are (exponentially) unstable by linear theory, but which ultimately approach a finite amplitude equilibrium. According to the general non-linear analysis, when two such disturbances interact it is possible that one of them decays to zero while the other approaches an equilibrium value. Which of the two disturbances decays in spite of being linearly unstable may be completely determined by the parameters of the problem, or may also depend on the initial amplitudes of the two disturbances. The only other possible equilibrium state for the model considered is composed of both disturbances, but this 'mixed' state will not occur if either of the interacting disturbances can ultimately decay to zero for some initial condition.

A mixed state may occur even if one of the two interacting disturbances is linearly stable. The reason is that for certain values of the parameters the growth of the originally unstable disturbance alters the sign of the growth rate of the originally stable disturbance. As explained below, there is some evidence to support a conjecture that this type of interaction may play a role in transition to turbulence.

These general results are applied to the thermal convection problem at various values of the Rayleigh number. For the most important situation to which our analysis applies, when the actual Rayleigh number is slightly above the minimum critical Rayleigh number of linear theory, it is shown that the mixed equilibrium state cannot occur. (This provides a first theoretical model illustrating the experimental fact that the non-linear terms appear to select for amplification one of the continuum of linearly unstable disturbances.) When the differences between the actual Rayleigh number and the critical Rayleigh numbers for each of the two disturbances are not too far apart, the initial conditions must be used to determine which disturbance decays to zero.

2. Mathematical formulation

We shall use the following notation and dimensionless variables: d is the distance between two horizontal planes bounding a fluid of mean density ρ, g is the acceleration of gravity (taken to act vertically downwards) and α_0, ν and κ are the coefficients of thermal expansion, kinematic viscosity and thermal conductivity, respectively. The dimensionless horizontal co-ordinates x and y and vertical co-ordinate z refer to the length d; similarly the corresponding velocities (u, v, w), the temperature θ, and time t refer to the scales κ/d, $\kappa\nu/\alpha_0 g d^3$ and d^2/κ. As is usual, we employ the Boussinesq approximation. A careful justification of this practice can be found in Spiegel & Veronis (1960). We restrict our consideration to two-dimensional 'rolls' so $v = \partial/\partial y = 0$. After some manipulation, as in Malkus & Veronis (1958), the equations can be written

$$u_x + w_z = 0, \tag{2.1}$$

$$\mathrm{T}_t - \mathrm{T}_{zz} = -(\overline{wT})_z, \tag{2.2}$$

$$T_t - \Delta T = -w\mathrm{T}_z - uT_x - wT_z + (\overline{wT})_z, \tag{2.3}$$

$$(\partial/\partial t - \sigma\Delta)\,\Delta w - \sigma T_{xx} = (uu_x + wu_z)_{xz} - (uw_x + ww_z)_{xx}, \tag{2.4}$$

where $\Delta \equiv \partial^2/\partial x^2 + \partial^2/\partial y^2$, σ is the Prandtl number ν/κ, and subscripts denote partial differentiation. The horizontal bar indicates an average in x,

$$\bar{q} = \lim_{L\to\infty} \frac{1}{2L}\int_{-L}^{L} q(x,z,t)\,dx, \tag{2.5}$$

and the temperature θ has been split into a mean and a fluctuating part:

$$\theta = \mathrm{T} + T, \quad \text{where} \quad \overline{\mathrm{T}} = \mathrm{T},\ \overline{T} = 0, \quad \text{so} \quad \bar{\theta} = \mathrm{T}. \tag{2.6}$$

The boundary planes at $z = 0$ and $z = 1$ are considered to be free surfaces (so that the normal velocity and tangential stress vanish) and perfect conductors. This gives the boundary conditions

$$T = w = w_{zz} = 0 \quad \text{on} \quad z = 0, 1. \tag{2.7}$$

If the top surface is kept at the temperature T_c and the bottom at T_h, $T_h > T_c$, the equations admit the steady solution

$$u = w = T = 0, \quad \mathrm{T} \equiv \mathrm{T}^{(00)} = -\mathscr{R}z, \tag{2.8}$$

where $\mathscr{R}$ is the Rayleigh number $\beta_0\alpha_0 g d^4/(\nu\kappa)$ and β_0 denotes the linear temperature gradient $\beta_0 = (T_h - T_c)/d$. (2.8) is the undisturbed solution which we perturb. Since the boundaries are held at constant temperature, we complete the perturbation boundary conditions by

$$\mathrm{T} - \mathrm{T}^{(00)} = 0 \quad \text{at} \quad z = 0, 1. \tag{2.9}$$

We now make a harmonic analysis of the temperature and velocity fluctuations, writing

$$q = q_{10}\cos\pi\alpha x + q_{01}\cos\pi\beta x + q_{11}\cos\pi(\alpha+\beta)\,x + q_{1-1}\cos\pi(\alpha-\beta)\,x + q_{20}\cos 2\pi\alpha x + q_{02}\cos 2\pi\beta x + \ldots, \tag{2.10}$$

where q stands for w and T, and α and β characterize the two interacting disturbances. The w_{ij} and T_{ij} are functions of z and t only. A similar series for u follows from the series for w and the continuity equation (1.1)

$$-\pi u = \alpha^{-1}(w_{10})_z \sin \pi\alpha x + \beta^{-1}(w_{01})_z \sin \pi\beta x + \dots \quad (2.11)$$

(An equivalent formulation involves sine series for w and T.)

The series (2.10) and (2.11) arise as follows: If we contemplate an initial unstable disturbance composed of two different fundamental wave-numbers, λ_1 and λ_2 say, the primary interaction of the λ_1 and λ_2 components with themselves and with each other (from terms like uT_x and wu_z) gives rise in general to four new components of wave-numbers $2\lambda_1$, $2\lambda_2$, $\lambda_1+\lambda_2$, $\lambda_1-\lambda_2$, and to alterations in the mean motion. A secondary set of interactions reproduces the fundamentals and introduces 14 new components such as $\lambda_1+3\lambda_2$, a tertiary set introduces many more new components, etc. In the problem which we consider, disturbances are initially so small that no interactions whatever need be considered. The disturbance amplitudes then increase exponentially according to linear stability theory until certain secondary interaction terms become comparable to the group of linear terms. Higher-order interaction terms are here comparatively insignificant and will be seen to remain so, for the disturbance amplitudes approach constant values. It will thus turn out that we need never consider the terms indicated by ... in (2.10) and (2.11).

To continue with the formalism, we assume

$$q_{ij} = q_{ij}^{(ij)} A^i B^j + q_{ij}^{(i+2,j)} A^{i+2} B^j + q_{ij}^{(i,j+2)} A^i B^{j+2} + \dots, \quad (2.12)$$

$$\mathrm{T} = \mathrm{T}^{(00)} + \mathrm{T}^{(20)} A^2 + \mathrm{T}^{(02)} B^2 + \dots, \quad (2.13)$$

where q again stands for w or T, the $q_{ij}^{(mn)}$ and $\mathrm{T}^{(mn)}$ depend only on z, and A and B (whose physical meaning is discussed below) depend only on t. Four comments will clarify (2.12) and (2.13). (1) It will be helpful to remember that $q_{ij}^{(mn)}$ multiplies $A^m B^n \cos \pi(i\alpha + j\beta)$. (2) In order that the first few terms of (2.12) and (2.13) give a good approximation to the answer, we must limit ourselves to functions A and B of small maximum absolute value. If squares and products of A and B are disregarded altogether

$$w = A(t)\, w_{10}^{(10)}(z) \cos \pi\alpha x + B(t)\, w_{01}^{(01)}(z) \cos \pi\beta x,$$
$$T = A(t)\, T_{10}^{(10)}(z) \cos \pi\alpha x + B(t)\, T_{01}^{(01)}(z) \cos \pi\beta x,$$

so that one may think of $A(t)$ and $B(t)$ as registering the approximate change with time of the α and β fundamentals. (3) For this problem, one can show that the corrections to q_{ij} and T are higher in order than the basic approximation by even powers of A and of B (e.g. $q_{ij}^{(i+1,j)} = q_{ij}^{(i,j+1)} = 0$). (4) Expansions (2.12) and (2.13) are simply a formalization of the successive interactions point of view discussed under (2·11).

The method of Stuart (1958) and Watson (1960) also requires series expansions of $\dot{A}$ and $\dot{B}$ ($\dot{} \equiv d/dt$)

$$\left.\begin{aligned} \dot{A} &= a^{(10)}A + a^{(30)}A^3 + a^{(12)}AB^2 + \dots, \\ \dot{B} &= b^{(01)}B + b^{(21)}A^2B + b^{(03)}B^3 + \dots \end{aligned}\right\} \quad (2.14)$$

Since the explicitly written terms are all that it will prove necessary to consider, we often use a simpler notation for the constants $a^{(ij)}$ and $b^{(ij)}$, writing the equations for A and B as

$$\dot{A} = aA - A[a_1A^2 + a_2B^2] + \ldots, \quad \dot{B} = bB - B[b_1A^2 + b_2B^2] + \ldots. \tag{2.15a, b}$$

Choice of the expansions (2.14), the keystone of the Stuart–Watson method, is justified by the fact that it enables one to find a meaningful formal solution to the problem under consideration. The choice of (2.14) may be *explained* as follows. We wish to base our non-linear analysis on the linear analysis. Not only is this mathematically convenient, but it is suggested by experimental evidence: in situations of the type we are investigating the flow closely resembles that predicted by linear instability theory except that the disturbance amplitude does not continue to grow exponentially but levels off to a constant value. It is perhaps 'natural' to try an expansion of the form

$$A = (\text{linear theory } A) + (\text{small corrections}). \tag{2.16}$$

But (2.16) is inappropriate because the linear theory A, $\exp(at)$, grows exponentially so that the correction must become large if A is to remain small, as is required for convergence of (2.12) and (2.13). What result of linear theory, then, can be expected to have a small correction when the non-linear terms are considered? The coefficient a is what we are looking for. If the Rayleigh number is sufficiently close to its critical value, linear theory shows that a is initially small. Speaking roughly, we expect a to be given by some function of A which ultimately approaches zero, so that the amplitude factor $\exp(at)$ will approach a constant as required for an equilibrating disturbance. More precisely if $A = \exp(at)$ then $\dot{A}/A = a$ so we expect that in the course of time, $\dot{A}/A$ will change from the small value a to zero, and will therefore remain uniformly small. Given this idea, the exact form of (2.14) follows from close examination of the equations (see also Watson 1960).

We proceed with the calculations by substituting (2.12), (2.13), and (2.14) into (2.1) to (2.4). Using trigonometric identities we equate to zero the coefficients of $A^iB^j \cos\pi(m\alpha + n\beta)x$ and obtain an infinite set of ordinary differential equations for the functions (of z) $w_{ij}^{(mn)}$, $T_{ij}^{(mn)}$ and $\mathrm{T}^{(mn)}$. We write here the equations for A^iB^j to the third order. For simplicity we make the abbreviations

$$\left.\begin{aligned} q_{10}^{(10)} &= q_1, \quad q_{01}^{(01)} = q_2, \quad q_{20}^{(20)} = q_3, \quad q_{02}^{(02)} = q_4, \quad q_{11}^{(11)} = q_5, \\ q_{1-1}^{(11)} &= q_6, \quad q_{10}^{(30)} = q_7, \quad q_{10}^{(12)} = q_8, \quad q_{01}^{(21)} = q_9, \quad q_{01}^{(03)} = q_{10}, \end{aligned}\right\} \tag{2.17}$$

(where q again stands for T or w) and

$$\mathrm{T}^{(20)} = \mathrm{T}_1, \mathrm{T}^{(02)} = \mathrm{T}_2. \tag{2.18}$$

During the course of the calculation, however, the original symmetric and meaningful notation of (2.1)–(2.14) proved valuable in guarding against errors and in keeping in mind the origin and meaning of the various terms. With the further abbreviations

$$L(a,\alpha;n) = aT_n - D_\alpha(T_n) - \mathscr{R}w_n, \quad D_\alpha(T) = T'' - \pi^2\alpha^2T,$$
$$M(a,\alpha;n) = a\sigma^{-1}D_\alpha(w_n) - D_\alpha^2(w_n) + \pi^2\alpha^2T_n, \quad ' = d/dz,$$

we have the following 22 equations for the 22 unknowns of (2.17) and (2.18):

$$O(A\cos\pi\alpha x):\quad L(a,\alpha;1) = M(a,\alpha;1) = 0; \tag{2.19}$$

$$O(B\cos\pi\beta x):\quad L(b,\beta;2) = M(b,\beta;2) = 0; \tag{2.20}$$

$$O(A^2\cos 2\pi\alpha x):\quad \left.\begin{aligned} L(2a,2\alpha;3) &= -\tfrac{1}{2}(T_1'w_1 - T_1w_1'),\\ \sigma M(2a,2\alpha;3) &= 2(w_1'w_1'' - w_1'''w_1);\end{aligned}\right\} \tag{2.21}$$

$$O(B^2\cos 2\pi\beta x):\quad \left.\begin{aligned} L(2b,2\beta;4) &= -\tfrac{1}{2}(T_2'w_2 - T_2w_2'),\\ \sigma M(2b,2\beta;4) &= 2(w_2'w_2'' - w_2'''w_2);\end{aligned}\right\} \tag{2.22}$$

$$O[AB\cos\pi(\alpha+\beta)x]:\quad \left.\begin{aligned} L(a+b,\alpha+\beta;5) &= -\tfrac{1}{2}[T_1'w_2 + T_2'w_1 - (\beta/\alpha)T_2w_1'\\ &\qquad - (\alpha/\beta)T_1w_2'],\\ \sigma M(a+b,\alpha+\beta;5) &= [(\alpha+\beta)/2\alpha]\,[w_1'w_2'' - w_1'''w_2\\ &\qquad + \pi^2(\alpha^2-\beta^2)w_1'w_2]\\ &\quad + [(\alpha+\beta)/2\beta]\,[w_1'w_2''\\ &\qquad - w_1'''w_2 - \pi^2(\alpha^2-\beta^2)w_1w_2'];\end{aligned}\right\} \tag{2.23}$$

$$O[AB\cos\pi(\alpha-\beta)x]:\quad \left.\begin{aligned} L(a+b,\alpha-\beta;6) &= -\tfrac{1}{2}[T_1'w_2 + T_2'w_1 + (\beta/\alpha)T_2w_1'\\ &\qquad + (\alpha/\beta)T_1w_2'],\\ \sigma M(a+b,\alpha-\beta;6 &= [(\alpha-\beta)/2\alpha]\,[w_1'w_2'' - w_1'''w_2'\\ &\qquad + \pi^2(\alpha^2-\beta^2)w_1'w_2]\\ &\quad - [(\alpha-\beta)/2\beta]\,[w_1'w_2'' - w_1'''w_2\\ &\qquad - \pi^2(\alpha^2-\beta^2)w_1w_2'];\end{aligned}\right\} \tag{2.24}$$

$$O(A^2):\quad 2a\mathrm{T}_1 - \mathrm{T}_1'' = -\tfrac{1}{2}(w_1T_1)'; \tag{2.25}$$

$$O(B^2):\quad 2b\mathrm{T}_2 - \mathrm{T}_2'' = -\tfrac{1}{2}(w_2T_2)'; \tag{2.26}$$

$$O(A^3\cos\pi\alpha x):\quad \left.\begin{aligned} L(3a,\alpha;7) &= a_1T_1 - \mathrm{T}_1'w_1 - \tfrac{1}{2}(\tfrac{1}{2}T_1w_3' + 2T_3w_1' + T_3'w_1\\ &\qquad + T_1'w_3),\\ \sigma M(3a,\alpha;7 &= \varpi_1 D_\alpha(w_1) + \tfrac{1}{4}(w_1''w_3' - w_1w_3''')\\ &\quad + (3\alpha^2/4)w_1w_3' - \tfrac{1}{2}(w_1'w_3'' - w_1'''w_3)\\ &\quad + (3\alpha^2/2)w_1'w_3;\end{aligned}\right\} \tag{2.27}$$

$$O(AB^2\cos\pi\alpha x):\quad \left.\begin{aligned} L(a+2b,\alpha;8) &= a_2T_1 - \mathrm{T}_2'w_1 - \tfrac{1}{2}[\beta(\alpha+\beta)^{-1}T_2w_5'\\ &\quad - \beta(\alpha-\beta)^{-1}T_2w_6'\\ &\quad + (\alpha+\beta)\beta^{-1}T_5w_2' - (\alpha-\beta)\beta^{-1}\\ &\qquad \times T_6w_2' + T_5'w_2 + T_6'w_2 + T_2'w_5\\ &\quad + T_2'w_6],\\ \sigma M(a+2b,\alpha;8) &= a_2D_\alpha(w_1) + \tfrac{1}{2}\alpha(\alpha-\beta)^{-1}[w_2''w_6'\\ &\quad - w_2w_6''' + (\alpha^2 - 2\alpha\beta)w_2w_6']\\ &\quad + \tfrac{1}{2}\alpha(\alpha+\beta)^{-1}[w_2''w_5' - w_2w_5'''\\ &\quad + (\alpha^2 + 2\alpha\beta)w_2w_5']\\ &\quad + (\alpha/2\beta)[-w_2'w_5'' + w_2'''w_5\\ &\quad + (\alpha^2 + 2\alpha\beta)w_2'w_5 + w_2'w_6'' - w_2'''w_6\\ &\quad - (\alpha^2 - 2\alpha\beta)w_2'w_6];\end{aligned}\right\} \tag{2.28}$$

$$O(A^2B\cos\pi\beta x):\quad \left.\begin{aligned} L(2a+b,\beta;9) &= b_1T_2 - T_1'w_2 - \tfrac{1}{2}[\alpha(\alpha+\beta)^{-1}T_1w_5 \\ &\quad + \alpha(\alpha-\beta)^{-1}T_1w_6' \\ &\quad + (\alpha+\beta)\alpha^{-1}T_5w_1' + (\alpha-\beta)\alpha^{-1} \\ &\quad\quad \times T_6w_1' + T_5'w_1 + T_6'w_1 + T_1'w_5 \\ &\quad + T_1'w_6], \\ \sigma M(2a+b;\beta;9) &= b_1D_\beta(w_2) - \tfrac{1}{2}\beta(\alpha-\beta)^{-1}[w_1''w_6' \\ &\quad - w_1w_6''' + (\beta^2-2\alpha\beta)w_1w_6'], \\ &\quad + \tfrac{1}{2}\beta(\alpha+\beta)^{-1}[w_1''w_5' - w_1w_5''' \\ &\quad + (\beta^2+2\alpha\beta)w_1w_5'] \\ &\quad + (\beta/2\alpha)[-w_1'w_5'' + w_1'''w_5 \\ &\quad + (\beta^2+2\alpha\beta)w_1'w_5 + w_1'w_6'' \\ &\quad - w_1'''w_6 - (\beta^2-2\alpha\beta)w_1'w_6]; \end{aligned}\right\} \quad (2.29)$$

$$O(B^3\cos\pi\beta x):\quad \left.\begin{aligned} L(3b,\beta;10) &= b_2T_2 - T_2'w_2 - \tfrac{1}{2}(\tfrac{1}{2}T_2w_4' + 2T_4w_2' \\ &\quad + T_4'w_2 + T_2'w_4), \\ \sigma M(3b,\beta;10) &= b_2D_\beta(w_2) + \tfrac{1}{4}(w_2''w_4' - w_2w_4''') \\ &\quad + \tfrac{3}{4}\beta^2w_2w_4' - \tfrac{1}{2}(w_2'w_4'' - w_2'''w_4) \\ &\quad + \tfrac{3}{2}\beta^2w_2'w_4. \end{aligned}\right\} \quad (2.30)$$

From (2.7) and (2.9) the boundary conditions are

$$T_n = w_n = w_n'' = 0 \quad \text{at} \quad z = 0, 1 \quad (n = 1, \ldots, 10) \quad \mathrm{T}_1 = \mathrm{T}_2 = 0 \text{ at } z = 0,1. \quad (2.31)$$

To avoid writing factors of π^4, let us define a modified Rayleigh number

$$\mathscr{R}' = \pi^{-4}\mathscr{R}. \quad (2.32)$$

No confusion should result from calling both $\mathscr{R}'$ and $\mathscr{R}$ 'the Rayleigh number' in qualitative contexts.

The main problem of linear stability theory is to find the lowest values of the Rayleigh number permitting a non-trivial solution of (2.19) for $a = 0$, the case of neutral stability. From the standard work on linear theory (Pellew & Southwell 1940) Rayleigh's result for the neutral stability curve is given by

$$\mathscr{R}' = \alpha^{-2}(\alpha^2+1)^3.$$

This curve is plotted as a solid line in figure 1. Points above the curve correspond to growing disturbances, and for a given α the Rayleigh number dividing growing from decaying disturbances is called the *critical Rayleigh number* for that α. The minimum critical Rayleigh number is $\mathscr{R}' = 27/4$ for $\alpha^2 = \frac{1}{2}$. 'Critical' will always stand for 'critical according to linear theory.'

We consider the Rayleigh number and two different unstable disturbance wave-numbers, α and β, to be given. For our series to converge, α and β must be chosen so that both the corresponding critical Rayleigh numbers, $\mathscr{R}'_\alpha$ and $\mathscr{R}'_\beta$, are close to the given Rayleigh number (see (4.3) and 3.3)). This limits α and β to values such that $(\alpha, \mathscr{R}')$ and $(\beta, \mathscr{R}')$ fall above the critical curve but below a vertically translated curve like the dashed one shown in figure 1. A possible set of values for α, β and $\mathscr{R}'$ is indicated on the figure.

We have assumed the undisturbed state (2.8) to be one of no motion and a linear temperature gradient. Under ordinary experimental conditions, $\mathscr{R}'$ would therefore have to be chosen just a little greater than 27/4, its minimum critical value. When $\mathscr{R}' - 27/4$ is not small, our analysis is formally valid provided that α and β are restricted in the way described in the previous paragraph. For (2.8) to

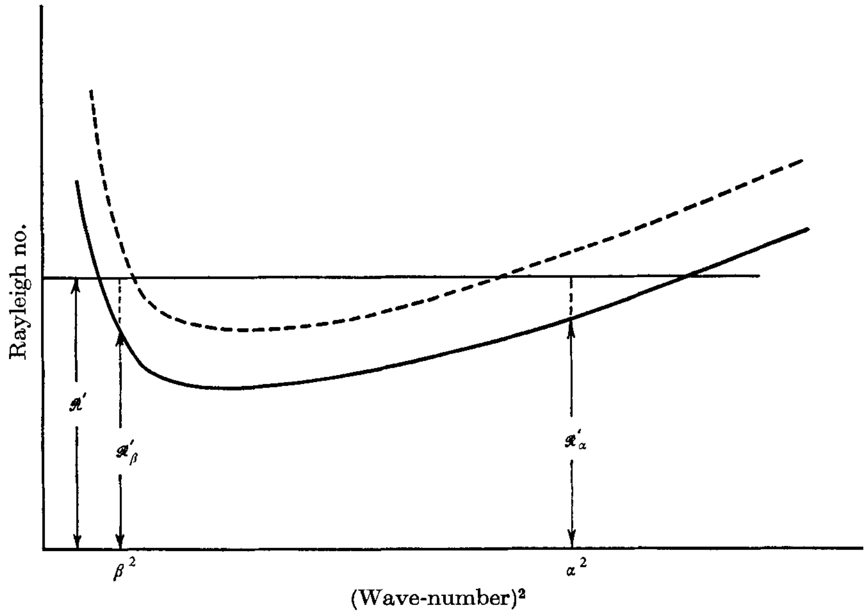

FIGURE 1. Values of α, β, and $\mathscr{R}$ for which the analysis is valid.

be the correct undisturbed state, however, even infinitesimal disturbances must be avoided until the α and β disturbances are excited. This is impossible in an actual physical situation, but might be achieved in a thought-experiment by using a conducting fluid and shutting off a stabilizing magnetic field at the same time that the two desired disturbances were artificially stimulated. We therefore shall not restrict our consideration to values of $\mathscr{R}'$ just above 27/4, the preceding remarks indicating that the results for larger values of $\mathscr{R}'$ might give information about possible qualitative behaviour.

3. Solution of the equations

We turn to the solution of equations (2.19)–(2.30). Solving (2.19) is the problem of linear stability theory. We have

$$\left.\begin{aligned} w_1 &= 2\sin\pi z, \quad T_1 = \pi^2\tau_1 \sin\pi z, \quad \tau_1 = 2\mathscr{R}'/(a'+\alpha^2+1), \\ w_2 &= 2\sin\pi z, \quad T_2 = \pi^2\tau_2 \sin\pi z, \quad \tau_2 = 2\mathscr{R}'/(b'+\beta^2+1), \end{aligned}\right\} \tag{3.1}$$

where $a' \equiv \pi^{-2}a$ is found from

$$\mathscr{R}' = \alpha^{-2}(\alpha^2+1)(a'+\alpha^2+1)(a'\sigma^{-1}+\alpha^2+1). \tag{3.2}$$

Since we assume $\mathscr{R}' - \mathscr{R}'_\alpha$ is small, it will be possible to ignore $(\mathscr{R}' - \mathscr{R}'_\alpha)^2$ and to obtain from (3.2)

$$\left.\begin{aligned} a' &= \sigma\alpha^2(\mathscr{R}' - \mathscr{R}'_\alpha)/(\sigma+1)(\alpha^2+1)^2 \\ \text{and similarly}\quad b' &= \sigma\beta^2(\mathscr{R}' - \mathscr{R}'_\beta)/(\sigma+1)(\beta^2+1)^2. \end{aligned}\right\} \tag{3.3}$$

We can also ignore a' in τ_1 and b' in τ_2. There are solutions to (2.19) and (2.20) proportional to $\sin n\pi z$, **n** a positive integer, but the solution with n = 1 has the lowest critical Rayleigh number. The factor multiplying $\sin \pi z$ in both w_1 and w_2 has for convenience been taken to be 2. Since w_1 and w_2 are multiplied by the still-to-be-determined quantities A and *B*, these factors may be chosen in any convenient way.

On substitution of the above expressions for w_1, w_2, T_1, and T_2, the right-hand sides of (2.21) and (2.22) reduce to zero. Since $\mathscr{R}'$ is nearly an eigenvalue for a and β, not 2α and 2β the only solution to these equations is

$$T_3 = T_4 = w_3 = w_4 = 0,$$

in agreement with the statement in Malkus & Veronis (1958) that there is no second-order distortion to the velocity field of a single roll. The right-hand sides of (2.23) and (2.24) do not reduce to zero, however, so that the two rolls interact to produce second-order fluctuations of wave-numbers equal to the sum and difference of the original wave-numbers. The amplitudes T_5, T_6, w_5 and w_6 turn out to be given by the formulas

$$\left.\begin{aligned} 2\alpha\beta D_1 T_5 &= \pi[(\beta-\alpha)(\alpha\tau_1 - \beta\tau_2)c_1^2 - 2\mathscr{R}'\sigma^{-1}(\beta^2-\alpha^2)^2]\sin 2\pi z, \\ 2\alpha\beta D_2 T_6 &= \pi[(\beta+\alpha)(\alpha\tau_1 + \beta\tau_2)c_2^2 - 2\mathscr{R}'\sigma^{-1}(\beta^2-\alpha^2)^2]\sin 2\pi z, \\ 2\alpha\beta D_1 w_5 &= \pi[(\beta-\alpha)(\alpha\tau_1 - \beta\tau_2)(\alpha+\beta)^2 - 2c_1\sigma^{-1}(\beta^2-\alpha^2)^2]\sin 2\pi z, \\ 2\alpha\beta D_2 w_6 &= \pi[(\beta+\alpha)(\alpha\tau_1 + \beta\tau_2)\,(a\text{-}/3)2 + 2c_2\sigma^{-1}(\beta^2-\alpha^2)^2]\sin 2\pi z, \\ \text{where}\quad c_1 &= (\alpha+\beta)^2+4,\quad c_2 = (a\text{-}/3)2+4,\quad D_1 = (\alpha+\beta)^2\mathscr{R}' - c_1^3, \\ D_2 &= (\alpha-\beta)^2\mathscr{R}' - c_2^3. \end{aligned}\right\} \tag{3.4}$$

The change in the mean temperature profile is easily found by solving (2.25) and (2.26). Using the fact that a and b are negligible compared to $2\pi^2$,

$$\mathrm{T}_1 = -\tfrac{1}{4}\pi\tau_2 \sin 2\pi z, \quad \mathrm{T}_2 = -\tfrac{1}{4}\pi\tau_2 \sin 2\pi z.$$

This completes the solution of the second-order equations.

We now find that everything on the right-hand side of the third-order equations (2.27)–(2.30) is known except for the four constants a_1, a_2, b_1, b_2—one in each equation-arising from (2.15). These constants can be determined uniquely in a manner carefully explained by Watson (1960). The general idea can be expressed as follows: we are dealing with problems of the type

$$L_a(y) + ky = F, \tag{3.5}$$

where L is a self-adjoint differential operator, k is a constant, F is a known function, and certain homogenous boundary conditions are to be imposed. *L, k* and F all depend on a small parameter a. For example, (2.27) has the form of (3.5). (In the present discussion, each of (2.19)–(2.30) will be considered an inhomogeneous equation for the appropriate w, T having been eliminated.) Let ϕ_n

denote the orthonormal set of eigenfunctions, and λ_n the corresponding eigenvalues, of the related problem

$$L_0(\phi) + \lambda\phi = 0, \tag{3.6}$$

where the subscript in L_0 denotes the fact that a has been set equal to zero in L. A standard way to solve (3.5) is to expand the known function F and the unknown function y in series of the ϕ_n's

$$Y = \Sigma c_n \phi_n, \quad F = \Sigma C_n \phi_n, \quad C_n = (F, \phi_n). \tag{3.7}$$

On substitution of (3.7) into (3.5) we find that if

$$c_n = C_n/(k-\lambda_n), \tag{3.8}$$

then y = $\Sigma c_n \phi_n + O(a)$, since $L_a(\phi_n) = -\lambda_n \phi_n + O(a)$. We now observe that, in the problem we are dealing with, k approaches one of the A,, *say* λ_N, as a $\to$ 0. (For example, as a $\to$ 0 the difference between (2.19) and the homogeneous part of (2.27) approaches zero, so λ_N represents the eigenvalue $\mathscr{R}$ of (2.19) with eigenfunction w_1.) We require that our solution remain finite as $a \to 0$. (This is the key point; for our problem it is clear that no singularities should occur as we approach the point of neutral stability, a = 0.) From (3.8), the finiteness requirement gives

$$\lim_{a\to 0} C_N = \lim_{a\to 0} (F, \phi_N) = 0, \tag{3.9}$$

or, for (2.27),

$$(\lim_{a\to 0} F, \lim_{a\to 0} w_1) = 0, \tag{3.10}$$

where (,) denotes the appropriate inner product

$$(f, g) = \int_0^1 f(z)\, g(z)\, dz.$$

The unknown constants a, and a_2 are thus determined to order one as required, by putting (2.27) and (2.28) into the form (3.5) and imposing (3.10). For b_1 and 6, we use (2.29) and (2.30) with w_2 instead of w_1 in (3.10). The approach given here can easily be generalized: for example, if L in *(3.5)* is not self-adjoint, we merely use the bi-orthogonality property to find C_N and replace ϕ_N in (3.9) by the function ϕ_N^+ satisfying the equation and boundary conditions adjoint to (3.6) when $\lambda = \lambda_N$.

Ignoring the small quantities a and b compared to unity (cf. (3.3) and the material preceding it), the values obtained for the constants are as follows:

$$a_1 = b_2 = \gamma, \quad \text{where} \quad \gamma = \tfrac{1}{2}\sigma/(1+\sigma), \tag{3.11}$$

$$a_2 = \gamma K\left[1 - \frac{Q_1(\beta/\alpha)}{2\sigma^2(\beta^2+1)^2}\right], \quad b_1 = \frac{\gamma}{K}\left[1 - \frac{Q_2(\alpha/\beta)}{2\sigma^2(\alpha^2+1)^2}\right], \tag{3.12}$$

$$K = (\alpha/\beta)^2[(\beta^2+1)/(\alpha^2+1)]^2, \tag{3.13}$$

$$Q_1 = W_6[\sigma^{-1}(\alpha^2+1)(3+\alpha^2-2\alpha\beta) - \tfrac{1}{2}\alpha\beta\tau_2] - W_5[\sigma^{-1}(\alpha^2+1)(3+\alpha^2+2\alpha\beta) + \tfrac{1}{2}\alpha\beta\tau_2] + \alpha(\tau_6-\tau_5), \tag{3.14}$$

$$Q_2 = -W_6[\sigma^{-1}(\beta^2+1)(3+\beta^2-2\alpha\beta) - \tfrac{1}{2}\alpha\beta\tau_1] + W_5[\sigma^{-1}(\beta^2+1)(3+\beta^2+2\alpha\beta) + \tfrac{1}{2}\alpha\beta\tau_2] - \beta(\tau_5+\tau_6), \tag{3.15}$$

$$W_5 \sin 2\pi z = [(\alpha-\beta)/(\alpha+\beta)]\, w_5, \quad W_6 \sin 2\pi z = [(\alpha+\beta)/(\alpha-\beta)]\, w_6, \tag{3.16}$$

$$\tau_5 \sin 2\pi z = (\alpha-\beta)\, T_5, \quad \tau_6 \sin 2\pi z = (\alpha+\beta)\, T_6.$$

For w_5, w_6, T_5 and T_6 see *(3.4);* for τ_1 and τ_2 see (3.1).

4. The nature of the amplitude functions

Since the right-hand aides are now completely known, we could proceed with the solution of (2.27)–(2.30), but we turn instead to the more informative investigation of the nature of the solutions to (2.15). For arbitrary constants c_1 and c_2, this pair of equations admits the special solutions

$$A = 0, \quad B = e^{bt}[c_2 + (b_2/b)e^{2bt}]^{-\frac{1}{2}}, \tag{4.1}$$

and

$$B \equiv 0, \quad A = e^{at}[c_1 + (a_1/a)\, e^{2at}]^{-\frac{1}{2}}, \tag{4.2}$$

which show how the non-linear terms can cause the equilibration of a disturbance which grows exponentially by linear theory. (In (4.1), for example, $B(t) \rightarrow (b/b_2)^{\frac{1}{2}}$ as $t \rightarrow$ co.) This phenomenon was first treated by Landau and is discussed in Landau & Lifschitz (1959, pp. 103–5). An independent investigation, and the first proceeding explicitly from the Navier-Stokes equations, was carried out by Stuart (1958).

We are here interested in the interaction of two disturbances so the special cases (4.1) and (4.2) are not of primary interest. It is not in general possible to solve (2.15) exactly, so we turn to an examination of the equilibrium points and their stability. From this we shall be able to obtain the desired qualitative behaviour of two interacting disturbances.

For the first part of the discussion, we shall consider (2.15) from a general point of view; later we shall see which of our results apply when the coefficients are determined by the formulas of the previous section. Setting $\dot{A} = \dot{B} = 0$ we obtain the following possible equilibrium points for (2.15):

$$\left.\begin{aligned} &\text{I}: && A = B = 0;\\ &\text{II}: && A = 0,\ B^2 = b/b_2;\\ &\text{III}: && B = 0,\ A^2 = a/a_1;\\ &\text{IV}: && A^2 = (a_2 b - a b_2)/(a_2 b_1 - a_1 b_2) \equiv \xi^2,\\ & && B^2 = (a b_1 - a_1 b)/(a_2 b_1 - a_1 b_2) \equiv \eta^2. \end{aligned}\right\} \tag{4.3}$$

The equilibrium value of $A(t)$ in III would occur if no B disturbance were present. As it should, the value obtained here, using (3.3) and (3.11), checks with the formula for ϵ in the corresponding part of the steady-state non-linear analysis of Malkus & Veronis (1958, pp. 235-5). The same check was obtained earlier in some unpublished work by J. Watson.

To examine the behaviour of solutions near the four groups of equilibrium points, we definenewvariables A_i, B_i; $i = 1, 2, 3, 4$; by translating the axes without rotation so that each equilibrium point in turn becomes the origin. Linearizing we obtain the following equations in the neighbourhood of the various equilibrium points

$$\text{I}: \quad \dot{A}_1 = aA_1, \ \dot{B}_1 = bB_1; \tag{4.4a}$$

$$\text{II}: \quad \dot{A}_2 = b_2^{-1}(ab_2 - a_2 b)\, A_2, \ \dot{B}_2 = -2bB_2; \tag{4.4b}$$

$$\text{III}: \quad \dot{A}_3 = -2aA_3, \quad \dot{B}_3 = -a_1^{-1}(ab_1 - a_1 b)\, B_3; \tag{4.4c}$$

$$\text{IV}: \quad \dot{A}_4 = -2a_1\xi^2 A_4 - 2a_2\xi\eta B_4, \ \dot{B}_4 = -2b_1\xi\eta A_4 - 2b_2\eta^2 B_4. \tag{4.4d}$$

For I, therefore, if either a or b is positive, $A = B = 0$ is an unstable equilibrium point. We will assume throughout this section that $\mathscr{R}$ is greater than both the critical values $\mathscr{R}_\alpha$ and $\mathscr{R}_\beta$ so that a and b are both positive—cf. (3.3)—and the origin is an unstable node. Turning to II, for these equilibrium points to exist b_2 must be positive, for if this were not the case, II in (4.3) shows that the equilibrium point would be imaginary. The same conclusion can be reached by observing that with $A = 0$, $b > 0$, and $b_2 < 0$, there are no stabilizing terms in (2.15). As higher-order terms would otherwise have to be considered we further limit ourselves in this section to cases where $b_2 > 0$. Hence II is a stable node if $ab_2 - a_2b$ is negative and a saddle (unstable) point if this quantity is positive. The same discussion holds *mutatis mutandis* for III. In particular, we assume $a_1 > 0$. For IV, a little more analysis is necessary, as in Andronow & Chaikin (1949, pp. 184–93). We find that IV is a stable node if $b_1a_2 - a_1b_2$ is negative and a saddle point if this quantity is positive. This result, like those preceding, holds no matter which of the two roots we take in solving for A and B in (4.3). We note here that standard theorems show that, for our problem, examination of the linear equations (4.4) gives the correct local behaviour of the original non-linear system. Also, due to the existence of the special solutions (4.1) and (4.2), the only possible limit cycle would be one encircling IV when it is a node. But, as we see below, if a node exists it is stable so that any limit cycle would be unstable and therefore of no real interest.

Suppose that both II and III are stable. We then have from (4.4)

$$ab_2 - a_2b < 0, \quad ab_1 - a_1b > 0, \quad \text{or} \quad a_2 > (a/b)\,b_2, \quad b_1 > (b/a)\,a_1. \tag{4.5}$$

On multiplying together the last two inequalities (the right-hand sides of which are positive) we obtain

$$a_2b_1 > b_2a_1. \tag{4.6}$$

Equations (4.5) and (4.6) show that in (4.3) ξ^2 and η^2 are indeed positive so that IV exists, but (4.6) also implies from the stability analysis just given that IV is a saddle point. The situation is illustrated in figure 2. One-quarter of the (A, B)-plane is shown; the remainder can be obtained by reflexion in the A and B axes. At each equilibrium point I, II and III, one trajectory (solution curve) approaches vertically and all others horizontally, or vice versa. What happens in a particular case can easily be worked out, as in Andronow & Chaikin (1949, p. 184–93). This type of analysis also shows that the slope of the line along which a trajectory approaches the saddle point IV is positive for all permissible values of the coefficients. We see that trajectories starting near the origin approach either II or III, depending on the initial mixture of A and B near I. The dividing case is the unstable trajectory leaving I and approaching the saddle point IV.

We have discussed the case when both II and III are stable, and the other possible cases are easier to understand. Altogether, only the following situations can occur when a, b, a_1, and b_2 are positive:

(1) II unstable, III unstable, if IV exists it is stable;
(2) II stable, III stable, IV cannot exist;
(3) II unstable, III stable, IV cannot exist;
(4) II stable, III stable, IV exists but is unstable.

Let us call an equilibrium solution containing only one wavelength and its harmonics a *pure state*; if other wavelengths are present we will speak of a *mixed state*. We may then summarize the whole situation as follows. We consider a class of disturbances each of which separately grows exponentially when infinitesimal but then approaches a finite-amplitude equilibrium. (Equivalently, we consider

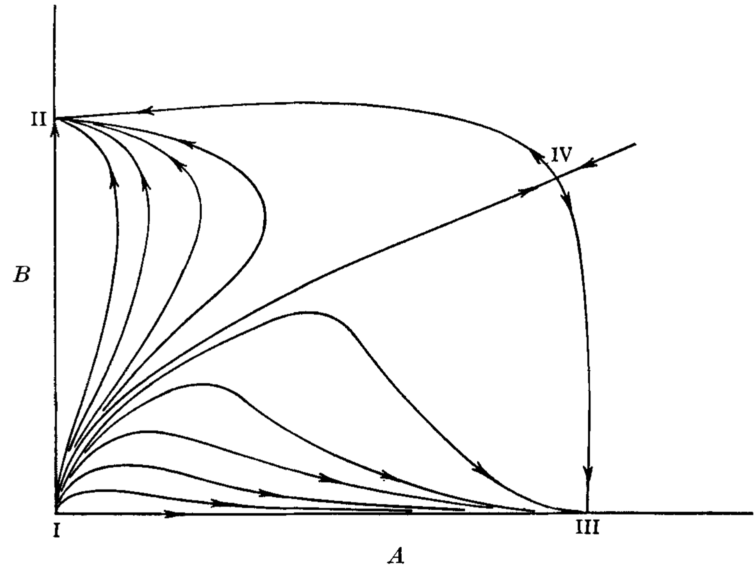

FIGURE 2. The interaction of two linearly unstable disturbances in which one disturbance decays to zero.

a, b, a_1 and b_2 to be positive.) When two such disturbances interact, *if a pure state can occur then a mixed state cannot occur.* (A state cannot occur physically unless it is stable.) Experimentally, when the Rayleigh number is above the critical and a whole band of wavelengths is unstable, the equilibrium state, nevertheless, seems to be very nearly pure. To show exactly how the non-linear terms bring this about does not yet seem feasible, but the above result gives a first indication of how the non-linear terms can single out a single unstable wavelength for equilibration.

5. Further general discussion of the amplitude functions

We have been discussing the possible fate of two disturbances which are linearly unstable, i.e. unstable by linearized theory. Let us consider for a moment whether the non-linear theory shows anything of interest when the two disturbances are linearly stable. In this case a and b are negative; we still assume that a_1 and b_2 are positive. The origin (I) is then a stable equilibrium point, while—from (19.1)—II and III cannot exist. To study IV, let us write $a = -c$ and $b = -d$ where b and d are positive. Now if IV is stable, we must have

$$b_1 a_2 - a_1 b_2 < 0. \tag{5.1}$$

If IV is to exist, however, this means—from (19.1)—

$$b_2 c < a_2 d, \quad d a_1 < c b_1. \tag{5.2}$$

All terms in (5.2) must be positive and we may multiply the inequalities and then divide both sides by cd, which yields $b_2 a_1 < a_2 b_1$, in contradiction to (5.1). We have therefore shown that non-linear theory verifies the linear prediction that the two disturbances considered will decay to zero.

In contrast to this, we find an interesting contradiction to linear theory by considering the interaction of a linearly stable disturbance with a linearly unstable disturbance. Suppose $b > 0$ and $a < 0$. (Modifications for $b < 0$, $a > 0$ are obvious.) Then III cannot exist (A alone cannot prevail) as we would guess.

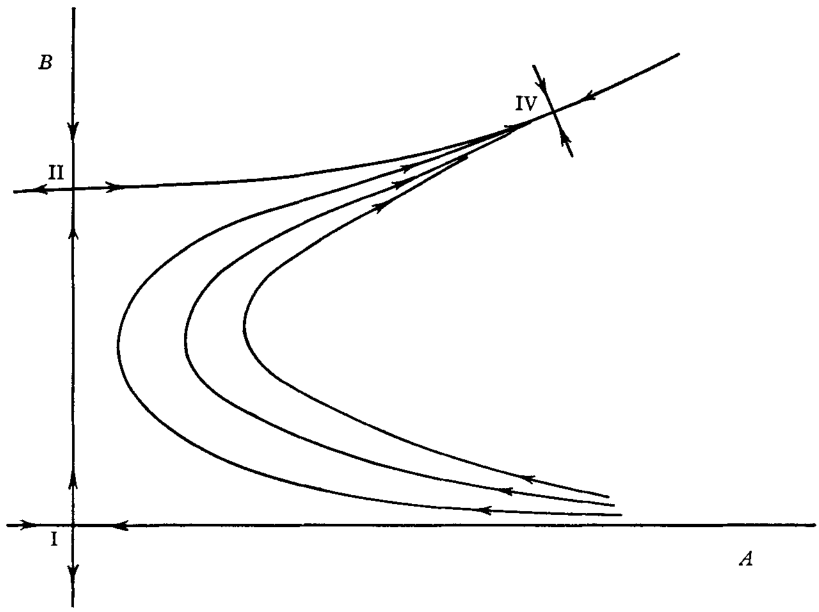

FIGURE 3. A linearly stable disturbance brought into a mixed equilibrium state on interaction with a linearly unstable disturbance.

From (4.3), II exists but, from (4.4), it will not be stable if $ab_2 - a_2 b > 0$. (For this it is clearly necessary that a_2 be negative.) Now if II is unstable (II will then be a saddle point) it is perfectly possible for IV to exist and be a stable node. The situation is depicted in figure 3. A pure A disturbance decays and a pure B disturbance grows to finite amplitude, but a combination of both will approach a mixed state containing a portion of linearly stable A as well as a portion of linearly unstable B. What happens is that although A at first decays, the growth of B changes the growth rate of A, ultimately forcing A to grow again. With this in mind, a glance at (2.15) shows again why it is necessary that a_2 be negative.

The situation just described may have more than academic interest. Lin (1955, p. 138) has pointed out that the behaviour of *damped* solutions to the linearized instability equations for parallel flows 'closely resembles the structure of turbulence' in that both high oscillatory and slowly varying regions are simultaneously present. This resemblance may be coincidental, but it is possible that in certain circumstances a non-linear mechanism like that just illustrated would, in the course of transition to turbulence, cause these initially damped disturbances to grow to a significant amplitude.

Further clarification of how the changing amplitude of one disturbance can affect the growth rate of another is found by looking for an approximate solution of (2.15) with $|B| \ll |A|$. We neglect the B^2 terms compared to the A^2 terms, which uncouples the equations. The solution turns out to be

$$A = e^{at}[C_1 + (a_1/a)\, e^{2at}]^{-\frac{1}{2}}, \quad B = C_2 e^{bt}[C_1 + (a_1/a)\, e^{2at}]^{-(b/2a_1)},$$

where C_1 and C_2 are arbitrary constants. In this approximation, then, A is unaffected by B. B grows like $\exp(bt)$ at first but later, under the influence of A, behaves like $\exp\{(b-ab_1/a_1)\,t\}$ in agreement with (4.4c).

In concluding this section, we note that when a_1 and b_2 are negative none of the equilibrium points II, III and IV are stable. If a and b are negative, all disturbances A and B of sufficiently small magnitude decay to $A = B = 0$ (so do large disturbances near $A = \pm B$ if II and III are both saddle points and IV does not exist.) In all other situations A and B grow without bound according to (2.15), so higher-order terms must be taken into account.

6. The amplitude functions in the present case

We now apply the general results of the previous two sections to the situation under investigation here. The formulas for a_1, a_2, b_1, and b_2 have been given, so that in any particular case we can in principle determine exactly what will happen. We have $a_1 = b_2 = \gamma$, which means that a single disturbance definitely equilibrates as our analysis requires. The expressions for a_2 and b_1 are so compliated, however, that it is best to consider some special cases.

Case 1: $\alpha - \beta$ positive and small. This would occur for any two unstable waves when $\mathscr{R}$ is only slightly greater than $\mathscr{R}_c$. In the present context, as explained at the end of § 2, this is the most important case physically. Here $a_2 = b_1 = 2\gamma$, approximately, when we neglect $\alpha - \beta$ compared to 1. We can now state the stability results in terms of a quantity μ,

$$\mu = (\mathscr{R} - \mathscr{R}_\alpha)/(\mathscr{R} - \mathscr{R}_\beta) = a/b. \tag{5.3}$$

For $0 < \mu < 2$, II is stable. For $\mu > \frac{1}{2}$, III is stable. For $\frac{1}{2} < \mu < 2$, both II and III are stable so that we have the case of figure 2. For all positive μ, either II or III is stable so IV cannot occur.

Case 2: $\beta \ll \alpha$ and $\beta^2 \ll 1$. This begins to be a good approximation when the Rayleigh number is twice the critical value. Taking advantage of the approximate equality of $\alpha^{-2}(\alpha^2+1)^3$ and $\beta^{-2}(\beta^2+1)^3 \approx \beta^{-2}$ to eliminate β, we find

$$b_1 = \gamma K^{-1}(1+q_1), \quad a_2 = \gamma K(1+q_2),$$

$$[(\alpha^2+4)^3 - (\alpha^2+1)^3]\, q_1 = \alpha^2[(\alpha^2+1)^2 + (\alpha^2+4)^2] + [2\alpha^4 - 3(\alpha^2+1)^3]\,\sigma^{-1} - [3\alpha^2(\alpha^2+1)(\alpha^2+4)]\,\sigma^{-2},$$

$$[(\alpha^2+4)^3 - (\alpha^2+1)^3]\, q_2 = (\alpha^2+1)^2[(\alpha^2+4)^2 - \alpha^2] + (\alpha^2+1)^3(2\alpha^2+3)\,\sigma^{-1} + \alpha^2(\alpha^2+1)(\alpha^2+3)(\alpha^2+4)\,\sigma^{-2}.$$

Consequently II, III, and IV are stable if $b(\mu - 1 - q_2)$, $b(1 - \mu - \mu q_1)$ and $q_1 + q_2 + q_1 q_2$

are respectively negative. The coefficient q_2 is always positive which favours the stability of II ($A = 0$, $B \neq 0$). On the other hand, q_1 changes from positive to negative as the Prandtl number σ decreases. This occurs for σ between 1 and 3, the exact value depending on α. A negative q_1 favours the instability of II and the appearance of B in conjunction with A even though B is stable.

Case 3: $\beta^2 = \frac{1}{4}$, $\alpha^2 = 1$. This is an intermediate case with values of α and β appropriate for $\mathscr{R}' \approx 8$. The qualitiative conclusions turn out to be about the same as case 2.

It is not prudent to draw refined conclusions from our idealized model (see the concluding remarks in Segel & Stuart 1962) and, as discussed above, special caution is necessary when $\mathscr{R} - (\mathscr{R}_c)_{\min}$ is not small. It may be significant, however, that since q_2 is always positive but q_1 is not, a shorter wavelength disturbance appears less apt to prevail than a longer. This probably reflects the fact that more energy is dissipated at shorter wavelengths. Also, it is possible but unlikely that mixed equilibrium states can occur, perhaps even as the result of a linearly stable disturbance growing under the influence of a linearly unstable disturbance. If mixed equilibrium states do occur the Rayleigh number will be somewhat above the critical value and the Prandtl number will probably be small. As explained in the Appendix the appearance of a mixed equilibrium state may be closely linked with transition to turbulence. It is therefore conjectured that heat-transfer experiments with fluids of small Prandtl number, like mercury, might give unusual results.

Our most compelling conclusions for the thermal problem are for the situation when the Rayleigh number is only slightly above its minimum critical value. One can say at once that the interaction of two linearly unstable roll disturbances cannot result in a mixed equilibrium state containing both disturbances. Further results are given in terms of μ as defined in (5.3). The quantity μ is the ratio of the differences between the actual Rayleigh number and the critical Rayleigh number for each of the two disturbances. Equivalently, neglecting the small difference between α and β, it is the ratio of the linear amplification rates.

When μ is not too far from unity, two equilibrium states are stable so the state finally attained is determined by the initial amplitude ratio of the two disturbances. (Two stable states were also found by Segel & Stuart 1962 in a related problem. In both situations, the stability of the equilibrium states to all possible disturbances remains an open question.) If the difference between the linear amplification rates is sufficiently great, however, the roll which grows faster according to linear theory will attain a finite amplitude and the other roll, while unstable by linear theory will ultimately decay to zero.

During the author's stay at the National Physical Laboratory he indirectly acquired an extensive training in non-linear stability theory from many conversations with J. T. Stuart and J. Watson. He is happy to record his gratitude for their considerable general assistance as well as for their specific suggestions concerning this paper. Acknowledgement is due to S. Tsao for drawing the figures. Finally, the author is grateful to the Office of Naval Research for encouragement and financial support.

Appendix: An alternative to Landau's successive instability theory of transition

Transition to turbulence is the process by which a fluid flow of determinate character changes into a flow which must be described statistically. In recalling Landau's theory of transition (Landau & Lifschitz 1959, pp. 103–7), let us consider a basic laminar flow characterized by a dimensionless number N. As N increases, a critical value will be reached and a disturbance to the basic flow will begin to grow. Landau conjectures that the mean-squared amplitude of this disturbance will level off at some finite value, thereby giving rise to a new basic flow. The process is repeated at a succession of critical values of N. Since it is not possible to determine precisely when an instability is triggered, each new basic flow contains a new arbitrariness in phase. We therefore have a possible transition mechanism, because if the phase relationships among many temporally periodic components of a flow are unknown then only statistical information can be expected.

During transition, the flow changes from one describable by one or two length scales to one whose description is best given by some sort of Fourier integral over a continuum of lengths. To explain how this might come about, we observe that the complete interaction of disturbances proportional to $\cos\lambda_1 x$ and $\cos\lambda_2 x$, say, gives terms proportional to $\cos\,(m\lambda_1+n\lambda_2)\,x$; $m, n = 0, \pm 1, \pm 2, \ldots$. Disregarding the possibility (of 'measure zero') that λ_1/λ_2 is rational, the quantity $m\lambda_1+n\lambda_2$ comes arbitrarily close to any given λ for appropriate choices of m and n, so it appears that successive instabilities need introduce only two 'rationally independent' length scales for each co-ordinate before a Fourier integral approach seems more appropriate than a Fourier series approach. It is possible, however, that interactions between different wavelengths become progressively weaker so rapidly that more different length scales are necessary before the Fourier spectrum is effectively filled in.

Let us consider the ideas just sketched in the light of what we found above concerning solutions to the amplitude equations (2.15a, b). Whenever conditions are such that a pure equilibrium state appears we have an instability of the Landau type, for if only one disturbance ultimately attains a finite equilibrium amplitude as each critical dimensionless number is exceeded, then at each stage one new phase arbitrariness is introduced.

On the other hand, the possible appearance of a mixed equilibrium state leads to an alternative transition mechanism. If, for example, we consider the interaction of 200 unstable disturbances and assume that a mixed state including all of these prevails, the approximation to a continuous spectrum afforded by even the first few interaction terms is a good one. Furthermore, since the interactions must now be described by a system of 200 differential equations, the solution will contain 200 arbitrary phase constants. An alternative to Landau's model of transition is therefore provided by the possibility that above some value of the relevant dimensionless number, many unstable infinitesimal disturbances to the basic flow will rapidly attain a finite amplitude. In order to be able to refer to it, we shall call this alternative the *multiple equilibrium theory*. The multiple equilibration might be connected with the instability of the original basic flow or of a flow obtained after several successive instabilities. Also, one ought to consider the interaction of the continuum of unstable disturbances, rather than the 200 mentioned above. On the other hand, if there should happen to be a good correlation between observed transition and the theoretical appearance of mixed equilibrium states upon the interaction of even two or three disturbances, we would feel that the multiple equilibrium theory provided a useful model for transition.

It should be mentioned that one must deal with three-dimensional flows before one can hope that either theory will account for any but the earliest phases of transition. In both theories, moreover, the equilibrium amplitudes of the disturbances considered must be periodic in time since initial phase differences lose their importance as steady equilibrium is approached.

Since stability calculations involving the simplest basic flows are difficult, it is hard to see how more than two or three successive instabilities could ever be calculated. In contrast, improving the multiple equilibrium model involves adding more disturbances to the *same* basic flow. If series methods like those used above are adequate or could be extended, calculations for an improved model should be of the same general type as for a simpler model. More algebra would be necessary to handle the increased number of interactions, but this might be handled by a computing machine. It is necessary to pick a basic flow, if one exists, whose instability leads to multiple equilibration. One possibility is the local flow preceding the appearance of a turbulent spot. Compared to the series of events in the Landau theory,the multiple equilibration of many disturbances to a single basic flow seems particularly well adapted to explain the very rapid turbulent burst.

REFERENCES

Andronow, A. & Chaikin, C. 1949 *Theory of Oscillations.* English edition: S. Lefschetz (ed.). Princeton University Press.

Landau, L. & Lifschitz, E. 1959 *Fluid Mechanics.* English Edition: trans. by S. Sykes & W. Reid. London: Pergamon Press.

Lin, C. C. 1955 *The Theory of Hydrodynamic Stability.* Cambridge University Press.

Malkus, W. V. R. & Veronis, G. 1958 *J. Fluid Mech.* **4**, 225.

Pellew, A. & Southwell, R. 1940 *Proc. Roy. Soc.* A, **176**, 312.

Segel, L. A. & Stuart, J. T. 1962 *J. Fluid Mech.* **13**, 289.

Spiegel, E. & Veronis, G. 1960 *Astrophys. J.* **131**, 442.

Stuart, J. T. 1958 *J. Fluid Mech.* **4**, 1.

Stuart, J. T. 1961 *Advances in Aeronautical Sciences, Vol.* 3/4, p. **121**. London: Pergamon Press.

Watson, J. 1960 *J. Fluid. Mech.* **9**, 371.