Chapter 10 The two –layer model

10.1 Formulation of the quasi-geostrophic equations for the two-layer model.

The complexity of the Charney model shows just how difficult it can be to examine the nature of the baroclinic stability problem when even the smallest step is made towards a realistic model, e.g. by including the β effect. Questions involving the role of friction, topography, variable stratification, or nonlinearity will be even more of a challenge. It is therefore important to find a model in which the basic process of baroclinic instability can take place and yet which, in the simplest case, is sufficiently easy so that we can confidently expect to make progress in situations where we include some of the physical process listed above. Such a model was introduced first by Phillips (1951, J. Meteor. **8**,381-394). It consists of two layers of fluid, each of constant density on the rotating beta plane. The interface between the two fluids is deformable and the entire baroclinicity of the system, i.e. the representation of a horizontal density gradient is represented by the slope of that interface. A detailed derivation of the quasi-geostrophic dynamics of this system can be found in chapter 6 (section 6.16) of GFD. Here I will present only a heuristic derivation relying on some facts you should be familiar with from earlier courses. The physical model is shown in figure 10.1.1
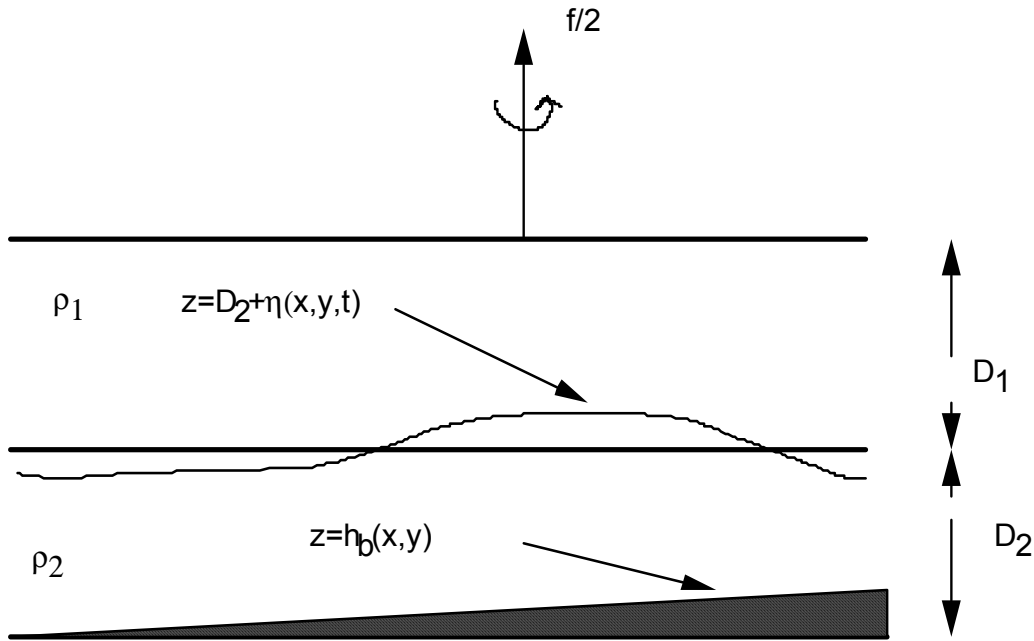
Figure 10.1.1 the two-layer model.

We will assume that the flow in each layer is geostrophic and hydrostatic except in frictional Ekman layers which may be considered at the lower solid boundary and the upper boundary if it is solid also. The external deformation is so much larger than the internal deformation radius based on the presumably small density difference, $\Delta \rho = \rho_2 - \rho_1$ that if the upper boundary is a free boundary it may be considered non deformable (it will produce negligible vortex stretching compared to the deviation of the interface). The undisturbed thickness of the upper layer is D_1. The thickness of the lower layer in the absence of motion (and so in the absence of a deviation of the interface) is $D_2 - h_b(x,y)$ and it is assumed that $h_b << D_2$. The density of each layer is a constant so the geostrophic velocity *within* each layer must be independent of the height coordinate, z. As in our discussion in chapter 2 the ratio of the vertical velocity to horizontal velocity is smaller than the simple geometrical scale estimate D/L by a factor of the Rossby number so that the advection operator lacks the vertical velocity term (and also because the horizontal velocity and vertical component of vorticity are independent of z

within each layer). Thus, within each layer the equation for ζ, the vertical component of vorticity is,

$$\frac{d}{dt}[\zeta_n + f] = f_o \frac{\partial w_n}{\partial z},$$

$$f = f_o + \beta y, \qquad \zeta_n = v_{n_x} - u_{n_y} \qquad\qquad (10.1.1\ a,b,c,d)$$

$$\frac{d}{dt} = \frac{\partial}{\partial t} + u_n \frac{\partial}{\partial x} + v_n \frac{\partial}{\partial y}$$

The subscript n refers to the layer. Upper layer variables are denoted by $n = 1$, lower layer variables by $n = 2$.

Integrate the vorticity equation over the depth of the upper layer to obtain,

$$(D_1 - \eta) \frac{d}{dt}[\zeta_1 + \beta y] = f_o[w_T - w_1(D_2 + \eta)] \qquad\qquad (10.1.2)$$

If the upper boundary were rigid the vertical velocity imposed there by the Ekman layer compatibility condition would be

$$w_T = -\frac{1}{2}\left(\frac{2A_v}{f_o}\right)^{1/2} \zeta_1 + w_s \qquad\qquad (10.1.3)$$

The first term on the right hand side is the vertical velocity pumped down from the upper Ekman layer assuming the boundary is solid. It is proportional to the upper layer vorticity. The second term w_s is a representation of an applied forcing. If the upper surface were solid it might be due to a differential rotation of the upper surface. If the upper surface were a free surface the first term would be absent and the second term might be due to an applied wind stress curl. Context will determine the combination of terms we will accept.

The kinematic condition at the interface yields,

$$w(D_2 + \eta) = \frac{d\eta}{dt} \tag{10.1.4}$$

It is straight forward to show by a simple scaling that the ratio of the interface deformation to the undisturbed layer thickness will be of the order of

$$\frac{\eta}{D_1} = O\left(\frac{f_o UL}{g'D_1}\right) = \varepsilon F_1,$$

$$\varepsilon = \frac{U}{f_o L}, \qquad F_n = \frac{f_o^2 L^2}{g'D_n} \tag{10.1.5 a,b,c}$$

The parameter F_n is a key parameter for the stability problem and is the ratio of the length scale L of the motion to the layer's deformation radius. The reduced gravity g' is just g multiplied by $\Delta\rho/\rho_o$ where ρ_O is the mean density of the two layers. With these approximations (10.1.2) becomes,

$$\frac{d}{dt}\left(\zeta_1 + \beta y + \frac{f_o}{D_1}\eta\right) = -\left(\frac{A_v f_o}{2D_1^2}\right)^{1/2}\zeta_1 + \frac{f_o}{D_1}w_s \tag{10.1.6a}$$

A similar integration over the depth of the lower layer yields,

$$\frac{d}{dt}\left(\zeta_2 + \beta y - \frac{f_o\eta}{D_2} + \frac{f_o h_b}{D_2}\right) = -\left(\frac{A_v f_o}{2D_2^2}\right)^{1/2}\zeta_2 \tag{10.1.6b}$$

In both (10.6.1 a and b) we recognize the terms in the brackets on the left as the quasi-geostrophic approximation to the layer representation of the potential vorticity.

The hydrostatic approximation implies that we can write the total pressure in each layer as,

$$p_{1_{total}} = -\rho_1 g(z - D_2) + p_1(x,y,t),$$

(10.1.7a,b)

$$p_{2_{total}} = -\rho_2 g(z - D_2) + p_2(x,y,t).$$

The total pressure must be continuous at the interface $z = D_2 + \eta$ so that,

$$\eta = \frac{p_2 - p_1}{\rho_o g'}$$

(10.1.8)

At the same time geostrophy implies

$$u_n = \frac{-1}{f_o} \frac{\partial p_n}{\partial y}, \quad v_n = \frac{1}{f_o} \frac{\partial p_n}{\partial x} \Rightarrow \zeta_n = \nabla^2 p_n / f_o$$

(10.1.9 a,b,c)

Defining the geostrophic streamfunction,

$$\psi_n = \frac{p_n}{\rho_o f_o}$$

(10.1.10)

allows us to (finally) write the quasi-geostrophic potential vorticity equation as,

$$\frac{d}{dt}\left[\nabla^2 \psi_n + \beta y + \frac{f_o^2}{g' D_n}(-1)^n (\psi_1 - \psi_2) + f_o \frac{h_b}{D_2} \delta_{n2}\right] = -\left[\frac{A_v f_o}{2 D_n^2}\right]^{1/2} \zeta_n + \frac{f_o}{D_1} w_s \delta_{n1}$$

(10.1.11)

Here we have used the kronecker delta function,

$$\delta_{ij} = 1, i = j$$
$$= 0, i \neq j$$

(10.1.12)

It will be convenient to introduce nondimensional variables. Let's choose U to be a scale for the horizontal velocity and L to be a scale for the horizontal coordinates x and y. Then (temporarily) denoting nondimensional variables by a prime, let,

$$\psi = UL\psi', \quad (x,y) = L(x',y'), \quad t = \frac{L}{U}t' \tag{10.1.13}$$

The quasi-geostrophic equations for the two layer model then become in nondimensional units (and where we have dropped primes on the nondimensional variables), and $n = 1,2$.

$$\frac{d}{dt}\left[\nabla^2\psi_n + \beta y + F_n(-1)^n(\psi_1 - \psi_2) + \eta_b\delta_{n2}\right] = -\frac{r_n}{2}\nabla^2\psi_n + S_n \tag{10.1.14}$$

$$F_n = \frac{f_o^2 L^2}{g' D_n}, \quad r_n = \frac{L}{U}\left(\frac{2A_v f_o}{D_n^2}\right)^{1/2}, \quad S_n = \frac{w_s L}{U D_1 \varepsilon}\delta_{n1}, \quad \eta_b = \frac{f_o h_b}{D_2}\frac{L}{U}$$

Note that,

$$\frac{d}{dt}\bullet = \frac{\partial}{\partial t}\bullet + J(\psi_n, \bullet) \tag{10.1.15.}$$

The constants r_n are the ratios of the advective time, L/U to the spin down time in each layer due to Ekman friction, $T_{n\,spin-down} = \frac{D_n}{[2A_v f_o]^{1/2}}$. The function η_b is the ratio of the change in the potential vorticity of the lower layer caused by topographic variations, compared to the characteristic relative vorticity and is analogous to the beta term in (10.1.5) which in dimensionless form is related to the dimensional beta by,

$$\beta = \frac{\beta_{\dim}L^2}{U} \tag{10.1.16}$$

which describes the change in planetary vorticity over the length L compared, again, to the relative vorticity. The potential vorticity in each layer is:

$$q_n = \nabla^2 \psi_n + (-1)^n F_n (\psi_1 - \psi_2) + \beta y + \eta_b \delta_{n2} \qquad (10.1.17)$$

The principal simplification that has occurred is that the system of equations contains only x, y and t as independent variables. The vertical coordinate z is no longer a continuous variable and is replaced by the index n which takes on only two values. As we shall see, the problem for the normal modes reduces to the solution of two coupled *ordinary* differential equations and, in the case where the basic flow is independent of y, it actually becomes an algebraic problem, a considerable simplification over the Charney model. It remains to see how effective the two –layer model is in recovering the essential features of that problem.

It is important to note that although the layer model can be thought of as a finite difference version of the continuous model, it is also the correct description of a self-consistent physical model consisting of n immiscible layers. Therefore no matter how few layers we consider the model is a correct one for the physical system described even when it is a poor representation of the continuous model. This important point implies that all the results to be described have physical validity even for the crudest model.

10.2 The formulation of the stability problem for the two-layer model

Consider now the total field described by (10.1.14) split between the basic flow and a perturbation,

$$\psi_n = \Psi_n + \phi_n,$$

$$\qquad (10.2.1 \text{ a,b})$$

$$q_n = Q_n + \tilde{q}_n$$

The basic state satisfies:

$$J(\Psi_n, Q_n) = -\frac{r_n}{2} \nabla^2 \Psi_n + S_n \qquad (10.2.2)$$

and if the source terms S_n are independent of the motion, the perturbation fields satisfy,

$$\frac{\partial \tilde{q}_n}{\partial t} + J(\Psi_n, \tilde{q}_n) + J(\phi_n, Q_n) + J(\phi_n, \tilde{q}_n) = -\frac{r_n}{2} \nabla^2 \phi_n \qquad (10.2.3)$$

We will, for the time being, consider the basic state to be a time-independent zonal flow, such that,

$$\Psi_n = \Psi_n(y), \quad \eta_b = \eta_b(y)$$

$$\qquad (10.2.4 \text{ a,b,c,d})$$

$$U_n(y) = -\frac{\partial \Psi_n}{\partial y}, \quad Q_n = \beta y + F_n(-1)^n [\Psi_1 - \Psi_2] + \frac{d^2 \Psi_n}{dy^2} + \eta_b \delta_{n2}$$

so that the equation for the perturbations can be written,

$$\left\{ \frac{\partial}{\partial t} + U_n \frac{\partial}{\partial x} \right\} \tilde{q}_n + \phi_{nx} Q_{ny} + J(\phi_n, \tilde{q}_n) = -\frac{r_n}{2} \nabla^2 \phi_n \qquad (10.2.5)$$

where

$$\tilde{q}_n = \nabla^2 \phi_n + (-1)^n F_n(\phi_1 - \phi_2),$$

$$\qquad (10.2.6 \text{ a,b})$$

$$Q_{ny} = \beta - (-1)^n F_n(U_1 - U_2) - U_{n_{yy}} + \eta_{b_y} \delta_{n2}$$

It would be useful at this point if you were to compare these equations with their continuous equivalents in chapter 3.

Along with the potential vorticity equation for the perturbations it is also useful, again, to examine the equation for the time rate of change of the x-averaged zonal flow. Repeating the same steps as in chapter 3, only now for the layer model (in nondimensional units) we obtain,

$$\frac{\partial \bar{u}_n}{\partial t} + \bar{v}_{n a} = -\frac{\partial}{\partial y} \overline{u_n' v_n'} - \frac{r_n}{2} \bar{u}_n \qquad (10.2.7)$$

The subscript a on v_n reminds us that this term is the ageostrophic velocity (it is the Coriolis torque term) while the prime variables on the right hand side forming the Reynolds stress terms are geostrophic and directly given by the geostrophic streamfunction. Once again a useful relationship between the potential vorticity flux and the Reynolds stress can be obtained (after several integrations by parts in x)

$$\overline{v'_n \tilde{q}_n} = \overline{v'_n \left\{ \left(v'_{n_x} - u'_{n_y} \right) + (-1)^n F_n (\phi_1 - \phi_2) \right\}} \qquad (10.2.8)$$

$$= -\frac{\partial}{\partial y} \overline{u'_n v'_n} - (-1)^n F_n \overline{\phi_n (\phi_1 - \phi_2)_x}$$

Let

$$d_n = D_n / D, \qquad D = D_1 + D_2 \qquad (10.2.9)$$

Multiply (10.2.8) by d_n and sum over n (equivalent to a vertical integration) to obtain,

(note that $d_n F_n$ is independent of n)

$$\sum_{n=1}^{2} \overline{v'_n \tilde{q}_n} d_n = -\sum_n \overline{(v'_n u'_n)_y} d_n + d_n F_n \overbrace{\frac{\partial}{\partial x} (\phi_1 - \phi_2)^2}^{=0} \qquad (10.2.10)$$

or,

$$\sum_{n=1}^{2} \overline{v'_n \tilde{q}_n} d_n = -\sum_n \overline{(v'_n u'_n)_y} d_n \qquad (10.2.11)$$

Using the fact that $\sum_n d_n \bar{v}_{n a} = 0$ by mass conservation, the equation for the rate of change of the zonal mean flow becomes,

$$\frac{\partial}{\partial t}\sum_n d_n \bar{u}_n + \sum_n r_n \bar{u}_n d_n /2 = \sum_n \overline{v'_n \tilde{q}_n} d_n \qquad (10.2.12)$$

so that the rate of change of the vertical integral of the x-averaged zonal flow is directly related to the potential vorticity flux of the perturbations. Since v'_n vanishes at the latitudinal end points of the region, the integral of the right hand side of (10.2.11) must vanish. Thus,

$$\int_{y1}^{y2} dy \sum_n \overline{v'_n \tilde{q}_n} d_n = 0 \qquad (10.2.13)$$

and this just states the obvious fact that in the absence of friction the total mean momentum in the x-direction can not change with time because the momentum fluxes can only redistribute that momentum and not change its total.

If we, for now, restrict our attention to small amplitude perturbations effectively linearizing (10..2.5) by ignoring the Jacobian term, the potential vorticity perturbation can be related to the small Lagrangian y displacement in each layer. As in chapter 3,

$$\tilde{q}_n = -\eta_n \frac{\partial Q_n}{\partial y} \qquad (10.2.14).$$

so that (10.2.12) becomes,

$$\frac{\partial}{\partial t}\sum_n d_n \bar{u}_n + \sum_n r_n \bar{u}_n d_n /2 = -\sum_n \frac{\partial}{\partial t}\frac{\overline{\eta_n^2}}{2}\frac{\partial Q_n}{\partial y} d_n \qquad (10.2.15)$$

which allows us to relate the forcing of the mean flow to the local dispersion of fluid elements in the y-direction in the basic state's potential vorticity gradient. Either from the point of view of momentum conservation (ignoring friction) or directly from (10.2.13) (which also assumes friction is unimportant)we obtain the condition.

$$\int dy \sum_n d_n \frac{1}{2}\frac{\partial}{\partial t}\overline{\eta_n^2}\frac{\partial Q_n}{\partial y} = 0 \qquad (10.2.16)$$

Thus if, as we would suppose, $\dfrac{\partial}{\partial t}\overline{\eta_n^2} > 0$ in each layer, it follows that for

(10.2.16) to be satisfied the basic state potential vorticity gradient must be both positive and negative in the meridional cross section of the basic current. It is *not* necessary that the gradient be zero anywhere. It may be, say, positive in the entire upper layer and negative in the lower layer and still satisfy (10.2.16). What we can say is that if the gradients are of the same sign over each layer the flow must be stable. In particular, a large enough value of β will stabilize the flow in distinction to the result in Charney's model and we will have to discuss why this is so. Clearly, we have restricted the vertical structure we are allowing to the perturbations and as a consequence the necessary condition for instability are stronger.

If we look for normal modes, i.e. if

$$\phi_n = \mathrm{Re}\,\Phi_n(y)e^{i\,k(x-ct)} \tag{10.2.17}$$

the ordinary differential equations governing the amplitude functions are,

$$(U_n - c)\left[\frac{d^2\Phi_n}{dy^2} - k^2\Phi_n + F_n(-1)^n(\Phi_1 - \Phi_2)\right] + \frac{\partial Q_n}{\partial y}\Phi_n = -\frac{r_n}{2ik}\left(\frac{d^2\Phi_n}{dy^2} - k^2\Phi_n\right) \tag{10.2.18}$$

where $\dfrac{\partial Q_n}{\partial y} = \beta - U_{n_{yy}} - (-1)^n[U_1 - U_2] + \eta_{b_y}\delta_{n2}$.

In common with the continuous model there exists a very similar semi-circle theorem and other bounds on the growth rate. Details can be found in chapter 7 of GFD.

10.3 The Phillips model of baroclinic instability.

In the reference to Phillips given above, he introduced the two-layer model's version of the Charney problem. Consider a zonal flow in which the basic velocity, U_n is independent of y. This implies, by (10.1.8) and (10.1.9 a) that the interface *slope* is a constant as shown in the figure below.
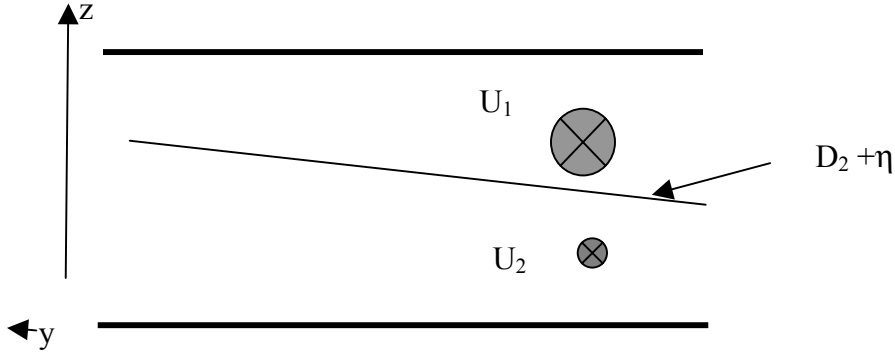
Figure 10.3.1 The interface slope in the two-layer model with a constant zonal flow in the upper layer and a smaller constant zonal flow in the lower layer. The interface slope is constant.

There is clearly potential energy in the sloping interface that can be released by the instability and we need to determine when that energy can be released and the structure of the perturbations responsible for the release. However, before executing the analysis of (10.2.18) recall that in the Charney model there is a characteristic vertical scale of the perturbation. When, as in the present model, the background density field is taken as constant that vertical scale is, (here the units are dimensional)

$$d_\beta = \frac{f_o^2}{N^2} \frac{U_{oz}}{\beta} \approx \frac{f_o^2 \Delta U}{g' \beta} \qquad (10.3.1)$$

where the last equality is a scaling approximation suitable to the layer representation in which

$$U_{oz} \approx U/D, \ N^2 = g\Delta\rho/\rho/D = g'/D \qquad (10.3.2)$$

Of course, should the scale given by (10.3.1) fall beneath the vertical scale resolved by the two layer model such perturbations could not be represented in the model. For weak shears Charney's model achieves instability by reducing the vertical scale of the perturbation so that its trajectories still lie within the wedge of instability. We must

anticipate that for weak shears in the two layer model, when $h_\beta < D_n$, the flow will appear stable since the model can no longer represent the surface intensified perturbations that are responsible for the instability. The condition that $h_\beta = D_n$ can then be written,

$$\Delta U = \frac{\beta_{\text{dim}}}{f_o^2 / g' D_n} \quad \text{(dimensional units)} \tag{10.3.3}$$

or in nondimensional units,

$$\Delta U_n = \frac{\beta}{F_n} \tag{10.3.4}$$

We consider the Phillips model with the following slight variations. Including the beta-effect and the possibility of a constant topographic slope in the y-direction, the potential vorticity gradients in the two layers are the constants,

$$Q_{1_y} = \beta + F_1(U_1 - U_2)$$

$$\tag{10.3.5 a,b}$$

$$Q_{2_y} = \beta - F_2(U_1 - U_2) + \eta_{by}$$

Recall, that the potential vorticity gradients must be both positive and negative for instability. Hence, since they are constants within each layer, one must be positive and the other negative for instability (at least in the absence of friction). If the flow has a positive shear, i.e. if $U_1 - U_2 > 0$ and if the bottom is flat, the necessary condition for instability tells us that

$$U_s = U_1 - U_2 > \beta / F_2 \tag{10.3.6}$$

for instability. We will have to check to see whether this *necessary* condition for instability is also *sufficient*, that is, whether the flow is really unstable as soon as the condition (10.3.6) is satisfied. Note that (10.3.6) has been anticipated by (10.3.4). The

flow is placed in a channel whose dimensional width is L so that using L as the scaling length the position of the boundaries are at $y = 0, 1$ where the perturbation meridional velocity must vanish.

Since the basic velocity and potential vorticity gradients are constant, the coefficients of (10.2.18) are independent of y. Therefore, solutions of the form,

$$\Phi_n = A_n \sin ly \qquad (10.3.7)$$

will satisfy the differential equation and the boundary conditions $y = 0, 1$ if $l = m\pi$ where m is an integer. Substituting (10.3.7) into the normal mode equation results in the following two *homogeneous* equations for the amplitudes,

$$A_1 \left| (c - U_1)(a^2 + F_1) + \left(\beta + \frac{ir_1 a^2}{2k} \right) + F_1(U_1 - U_2) \right| = A_2 (c - U_1) F_1$$

$$\qquad (10.3.8a,b)$$

$$A_2 \left[(c - U_2)(a^2 + F_2) + \left(\beta + \eta_{b_y} + \frac{ir_2 a^2}{2k} \right) - F_2(U_1 - U_2) \right] = A_1(c - U_2) F_2$$

where the total wave number is a, i.e.

$$a^2 = k^2 + l^2$$

Note that we have retained the effect of Ekman friction in the two layers.

Since the algebraic equations for the perturbation amplitudes are linear and homogeneous a non trivial solution requires the determinant of the coefficients to vanish. This condition will yield a quadratic equation for the phase speed c which we can consider a function of wavenumber a. This dispersion relation is generally complex because of the Ekman friction terms but even if they are ignored the phase speeds may themselves be complex and a positive imaginary part for c implies instability. It is useful to introduce the notation,

$$B_n = \beta + \frac{ir_n a^2}{2k} + \eta_{b_y} \delta_{n2} \qquad (10.3.9)$$

After considerable algebra the quadratic equation for c which comes from the condition for nontrivial solutions of (10.3.8 a,b) yields as solutions,

$$c = U_2 + \frac{U_s}{2} \frac{a^2 + 2F_2}{a^2 + F_1 + F_2} - \left\{ \frac{B_2(a^2 + F_1) + B_1(a^2 + F_2)}{2a^2(a^2 + F_1 + F_2)} \right\}$$

$$\pm \frac{\left\{ \begin{array}{l} \left[U_s^2 a^4 (a^4 - 4F_1F_2) + 2Usa^2 \left[(B_2 - B_1)(a^4 - 2F_1F_2) + a^2(B_2F_1 - B_1F_2) \right] + \right.^{1/2} \\[2mm] \left[B_2(a^2 + F_1) - B_1(a^2 + F_2) \right]^2 + 4F_1F_2B_1B_2 \end{array} \right\}}{2a^2(a^2 + F_1 + F_2)}$$

$$(10.3.10)$$

where the shear, as in (10.3.6) is defined as,

$$U_s \equiv U_1 - U_2.$$

Once c has been found, the ratio A_1 / A_2. (the vertical structure) can be found from either (10. 3. a or b). The algebraic solution given in (10.3.10) is a complicated one and we shall examine some familiar special cases but first note how easily, in the two layer model, we have achieved a solution for the stability problem with beta, topography and friction.

10. 4 Simple examples

   a) Frictional spin down

     Consider first the case in which there is no basic velocity, no beta, and no topographic slope and in which $F_1 = F_2 \equiv F$. When $F_1 = F_2$ (i.e. $H_1 = H_2$)this is equivalent to saying that the stratification in the continuous system is independent of z (why?). Note that this also implies that $r_1 = r_2 = r$. With these restrictions $B_n = ira^2/2k$ and the two roots for c are,

$$c = -i\frac{r}{2k} \left\{ 1, \frac{a^2}{a^2 + 2F} \right\} \qquad\qquad (10.4.1)$$

Both roots are decaying with decay rates,

$$kc_i = \frac{1}{2} \begin{cases} -ir \\ -ira^2/(a^2 + 2F) \end{cases}$$

(10.4.2a,b)

The first root $-ir/2$ is independent of wave length and, we shall see, represents the spin down of the barotropic mode. Indeed the time scale associated with this root is the classical spin down time of Greenspan and Howard (1963, J. Fluid Mech. **17**, 385-404). The second root yields the spin down rate of a purely baroclinic mode. It does depend on the wavelength. When $a^2 \gg 2F$, or equivalently when the length scale is small compared to a deformation radius, the spin down time is equal to the barotropic value. On the other hand, when the length scale is large compared to the deformation radius most of the potential vorticity is manifested in the stretching term (i.e. the layer thickness) and the elimination of the relative vorticity by the Ekman friction does little to alter the potential vorticity and so for small a the decay rate is small and goes to zero as a → 0. A little algebra shows that for the first root, (10.3.8 a) yields $A_1 = A_2$ so, as expected, the corresponding motion is barotropic. For the second root the same equation (for $H_1 = H_2$) yields $A_1 = -A_2$ so that the mode decaying with this root is purely baroclinic. You should check how an arbitrary initial condition decays as a combination of these two modes.

b) Rossby waves.

Now let's add the beta effect to the previous example, still keeping the shear zero and ignoring any effect of topography . We will also ignore topography but allow the two layer thickness to be different. In this case $B_n = \beta$ and the two roots for c are,

$$c_1 = -\frac{\beta}{a^2},$$

<div align="right">(10.4.3 a,b)</div>

$$c_2 = -\frac{\beta}{a^2 + F_1 + F_2}$$

The first root corresponds to the dispersion relation for barotropic Rossby waves. The second corresponds to the dispersion relation for baroclinic Rossby waves. If each root is substituted into (10.3.8a) we obtain for the first root, again, $A_1 = A_2$ (barotropic). The second root yields,

$$\frac{A_1}{F_1} = -\frac{A_2}{F_2} \qquad \Rightarrow A_1 D_1 = -A_2 D_2$$

or a purely baroclinic (no net transport) mode. You should check in the simpler case when $D_1 = D_2$ that the addition of friction would add to each of the roots of (10.4.3 a,b) the corresponding decay rates of (10.4.2 a,b).

c) The inviscid stability problem

Now let's consider the important case in which there is now a basic flow and a non zero vertical shear. Again we will ignore topography and temporarily Ekman friction. The solution for the phase speed reduces to,

$$c = U_2 + \frac{U_s(a^2 + 2F_2)}{2(a^2 + F_1 + F_2)} - \beta \frac{\left\{2a^2 + F_1 + F_2\right\}}{2a^2\left\{a^2 + F_1 + F_2\right\}}$$

<div align="right">(10.4.4)</div>

$$\pm \frac{\left\{U_s^2 a^4 \left[4F_1F_2 - a^4\right] + \beta^2(F_1 + F_2)^2 + 2a^4 U_s\beta(F_1 - F_2)\right\}^{1/2}}{2a^2\left\{a^2 + F_1 + F_2\right\}}$$

The radicand of (10.4.4) may become negative in which case the roots for c will be complex conjugates. This would yield instability. The critical condition where the

radicand just vanishes gives us the curve of marginal stability. This determines a critical shear from the condition that the radicand vanish or,

$$U_s{}^2 a^4 \left(4F_1F_2 - a^4\right) - 2a^4 U_s \beta (F_1 - F_2) - \beta^2 (F_1 + F_2)^2 = 0 \qquad (10.4.5)$$

Traditionally, one asks for the shear required to satisfy (10.4.5) for a given value of β as a function of wavenumber. The value of shear which satisfies (10.4.5) is,

$$U_s = \frac{\beta(F_1 - F_2)}{4F_1F_2 - a^4} \pm \frac{2\beta(F_1F_2)^{1/2}\left[(F_1 + F_2)^2 - a^4\right]^{1/2}}{a^2\left(4F_1F_2 - a^4\right)} \qquad (10.4.6)$$

If the layer depths are equal so that $F_1 = F_2 \equiv F$, the critical condition reduces to

$$U_s = \pm \frac{2\beta F}{a^2\left(4F^2 - a^4\right)^{1/2}} \qquad (10.4.7)$$
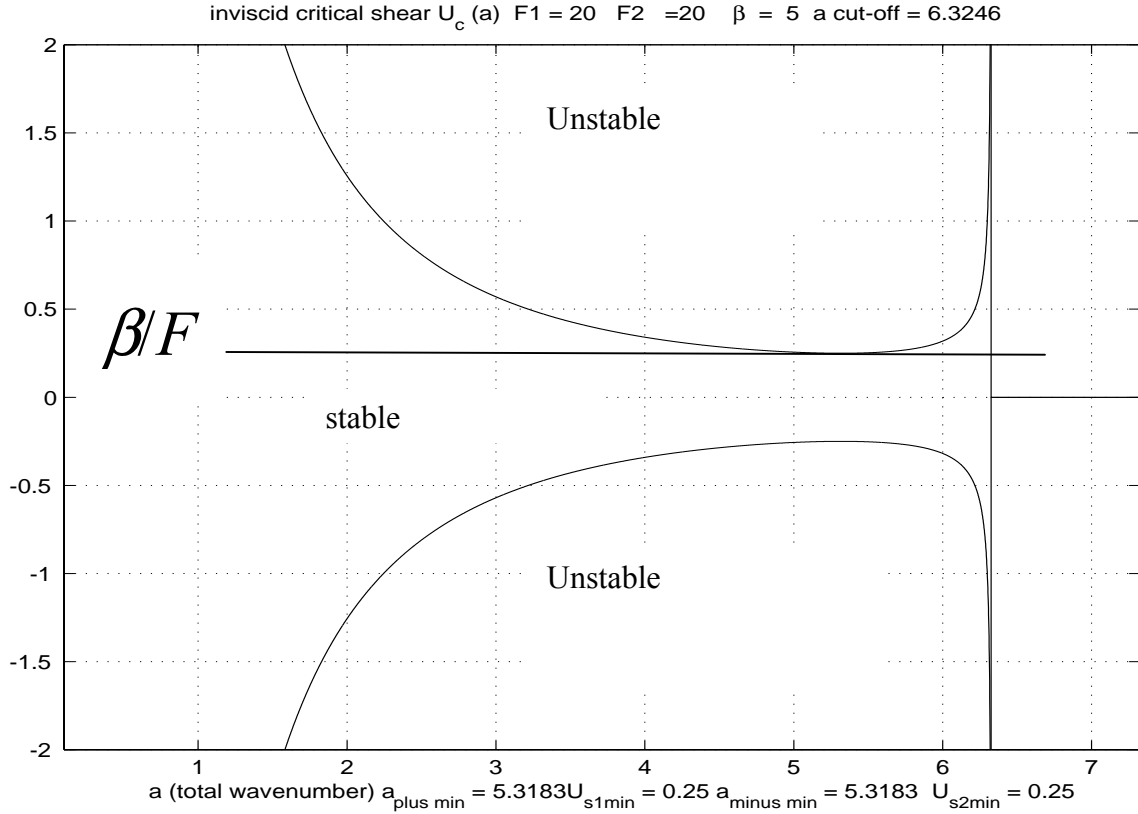
and is shown in Figure (10.4.1).

Figure 10.4.1 The marginal stability curves when both layer thicknesses are equal.

When the shear exceeds the value given by the curves in Figure 10.4.1 the radicand in (10.4.4) becomes negative and c becomes complex with a non zero imaginary part. There is, by direct calculation a minimum value of the critical shear,

$$U_{s\min} = \frac{\beta}{F} \quad \text{at } a^2 = \sqrt{2}F \tag{10.4.8}$$

It is important to note that in this case the necessary condition for instability, i.e. the condition that the potential vorticity gradient be of opposite signs in the two layers is also the *sufficient* condition for instability as established by direct calculation. For positive shear this occurs when the potential vorticity gradient in the lower layer,

$$Q_{2y} = \beta - FU_s \tag{10.4.9}$$

just vanishes. Note that there is a short wave cut-off also, That is, independent of β there is always stability if $a^2 > 2F$. This is qualitatively equivalent to the Eady short wave cut-off. Just as we anticipated , for weak enough shears, where the Charney model would yield vertical scales less than the layer depth, the two-layer model is stable. One could either consider this a disadvantage of the layer model, i.e. its inability to describe small scale, low shear modes with small vertical scales, or as an advantage in the sense that it filters out those small scale weakly growing modes.

The real and imaginary parts of c and the growth rate, kc_i are shown in Figure 10.4.2,
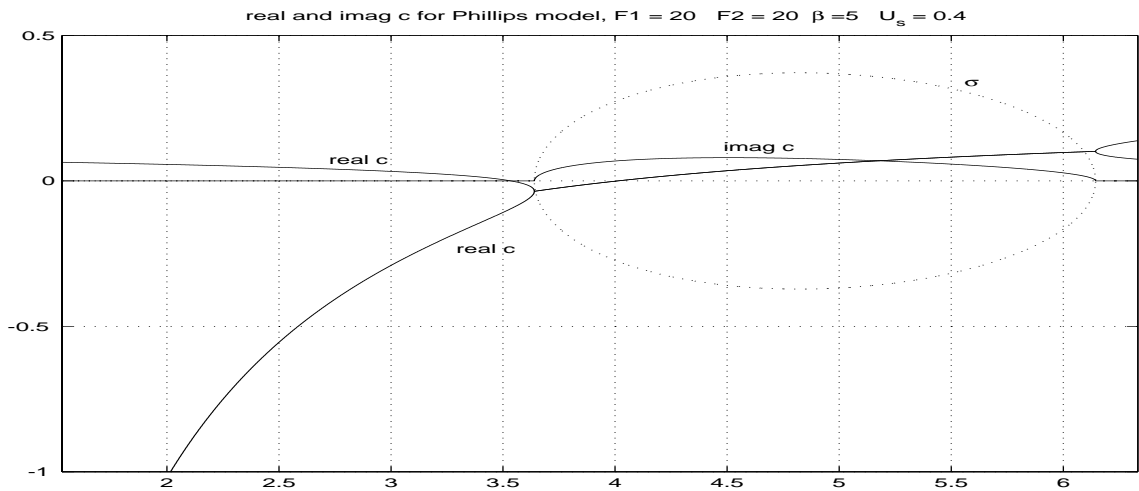

Figure 10.4.2 The real and imaginary parts of the phase speed and the growth rate σ.

In the figure we see the two real roots for c coalesce at the two boundaries of the unstable region inside of which the real part of the phase speed is given by the first term on the right hand side of (10.4.6) and the imaginary part is given by the square root term. The growth rate σ is given by kc_i. In the figure the variables are plotted against the total wave number a for the case in which $l=\pi/10$ at a value of the shear which is 0.4 whereas the minimum critical value is 0.25.

The situation becomes more complicated when the layer thickness are unequal. If we were to apply the model to the stability of an oceanic flow, like the Gulf Stream we might think of putting the current in a relatively thin upper layer and allowing the lower layer to

be considerably thicker, perhaps representing the ocean below the thermocline. The two potential vorticity gradients are then,

$$Q_{1_y} = \beta + F_1 U_s,$$

$$Q_{2_y} = \beta - F_2 U_s$$

If the shear is positive, i.e. if $U_1 > U_2$ the minimum value of the shear required for instability would be,

$$U_{s+} = \frac{\beta}{F_2} \qquad (10.4.10a)$$

while for negative shear the critical value is

$$U_{s-} = -\frac{\beta}{F_1} \qquad (10.4.10b)$$

If we suppose that $D_2 > D_1$ it follows that $F_2 < F_1$ so that a larger positive, eastward shear is required for instability than for the case of westward, negative shear since,

$$\frac{U_{s+}}{|U_{s-}|} = \frac{D_2}{D_1} > 1 \qquad (10.4.11).$$

The stability diagram is shown in figure 10.4.3 and, again, the necessary conditions for instability turn out to be sharp ones in the sense that they are also sufficient conditions.
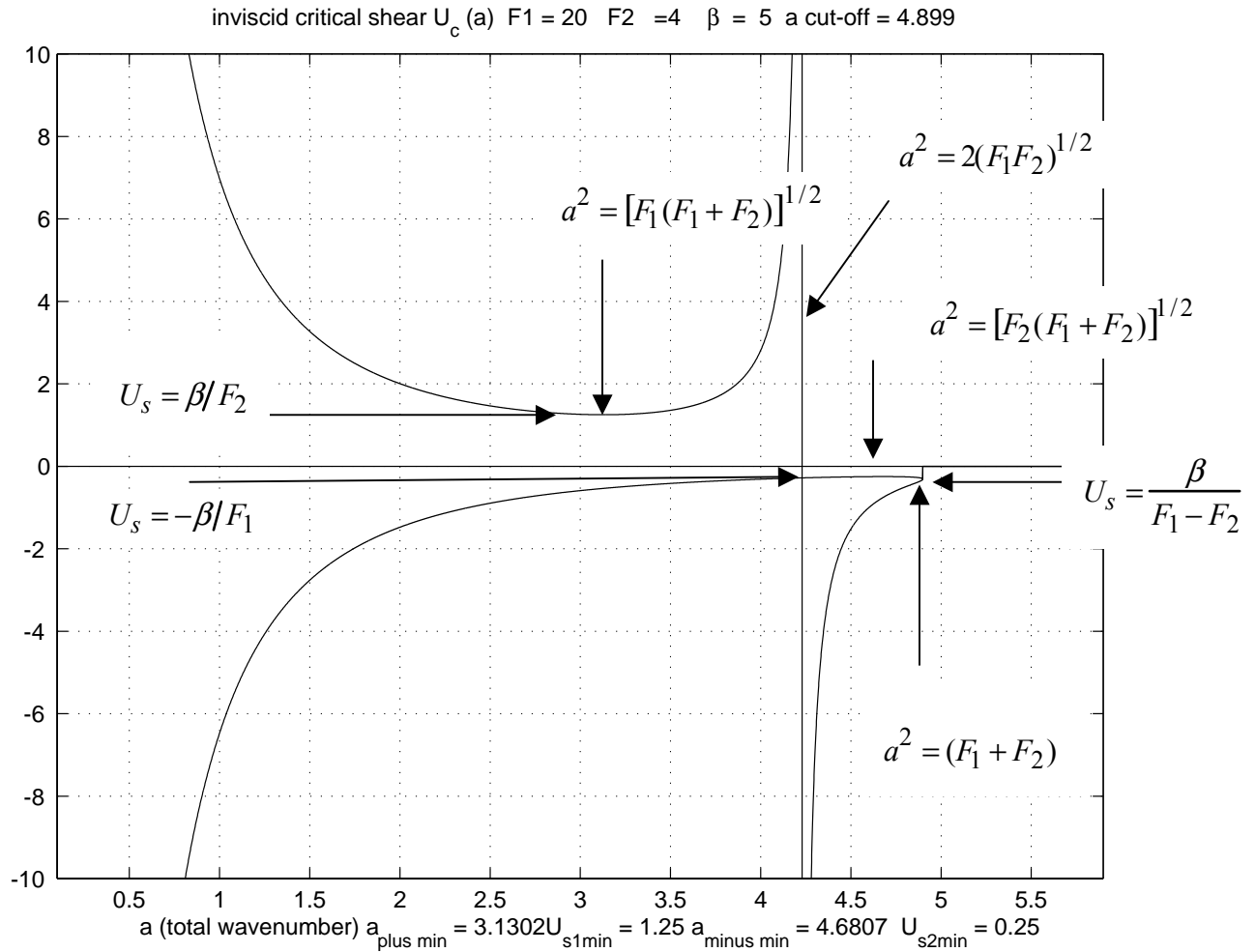
Figure 10.4.3 The marginal stability curves for unequal layer depths in the case $D_1 < D_2$.

In the figure particular wave numbers and shear levels are noted. Note that for the negative shear case not only is the minimum shear required for instability smaller, but the wavenumbers of unstable modes extend to much higher values. Indeed, in the limit as $D_1 \to 0$ the "nose" of the curve extends to very large wavenumber and to very small values of the shear. If the depth ratios were reversed, it would be the positive shear case that would have the small minimum value for the critical shear and be unstable at large wavenumber. In that case the result of the two layer model would increasingly resemble the critical curve for the Charney mode. Reducing D_2 allows us to represent the Charney mode for weak shear since for weak shear the perturbation was seen to be trapped to a thin region near the lower boundary.

The growth rate for the case of negative shear is shown in Figure 10.4.4a. A shear about 1 1/2 times critical has been chosen. Note that the growth rate peaks at a rather large wavenumber (small scale) compared to the case for positive shear which is shown in Figure 10.4.4 b.
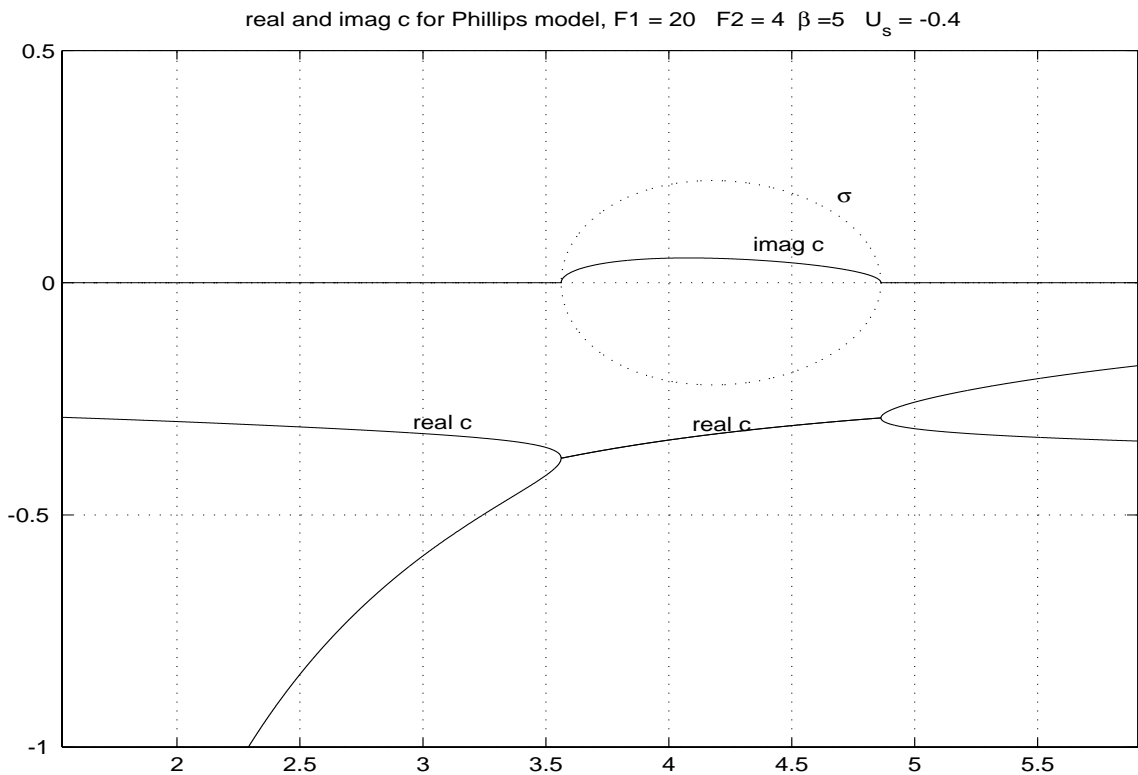

Figure 10.4.4 a The real and imaginary parts of c and the growth rate σ for the case of negative shear. The shear in nondimensional units is –0.4 whereas the minimum critical shear is –0.25. This should be compared with the case of positive shear shown in Figure 10.4.4b. There the minimum critical shear is 1.25 and the chosen shear, in the same proportion as in the negative shear example, is 2.0

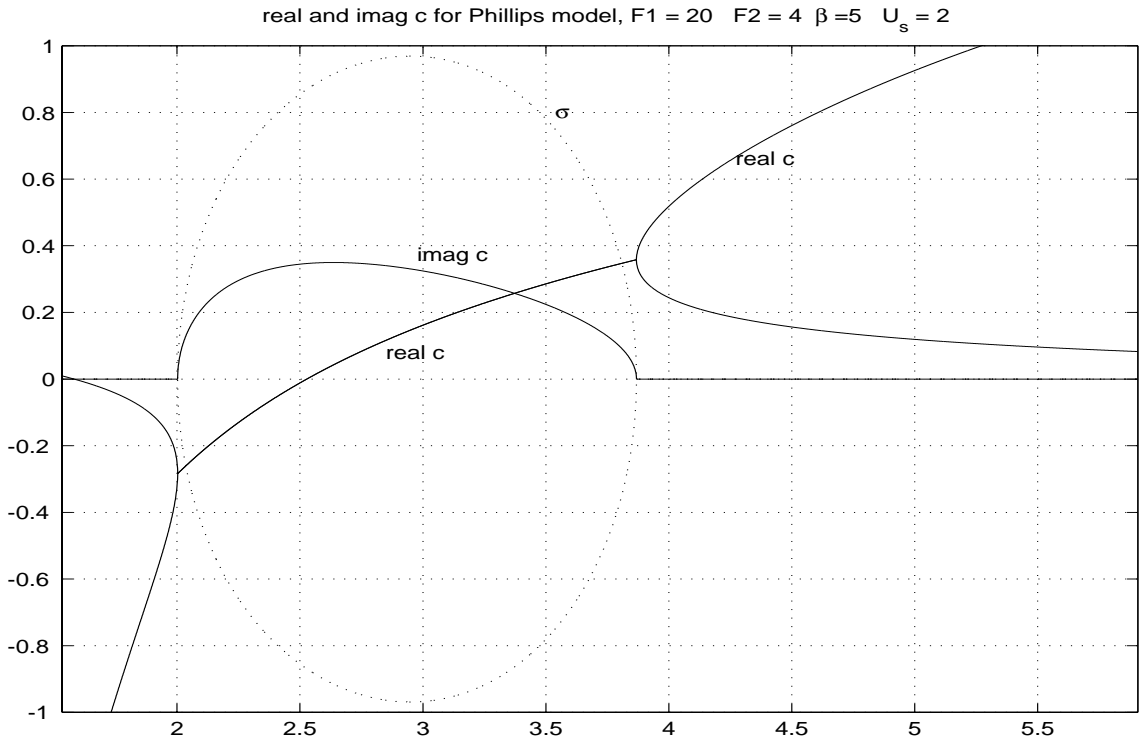real and imag c for Phillips model, F1 = 20   F2 = 4   β =5 U_s = 2

Figure 10.4.4b As in the previous figure except that the shear is positive. Note that the peak in the growth rate is at smaller wave number.

When the shear is set at the minimum critical shear both roots for *c* are real and they coalesce at the point of marginal stability on the bottom of the neutral curve. Figure 10.4.4c shows the case when, for the same parameters as figure 10.4.4a we choose the shear to be the critical value 1.25. We have (with a Galilean transformation) chosen the lower layer basic state velocity to be zero without loss of generality. Note that one root is *always c* = 0.The other root coalesces at that point at the wavenumber that is corresponds to the marginal curve $a = \{F_2[F_1 + F_2]\}^{1/4}$ which in the present case is about 3.13. This implies that the whole lower layer becomes a critical point (or layer) in that $U_2 - c = 0$ in the layer at the minimum critical shear on the marginal curve. Thus both linear terms in the potential vorticity equation identically vanish there and we must expect that nonlinearity or friction would be especially important for such parameter values.
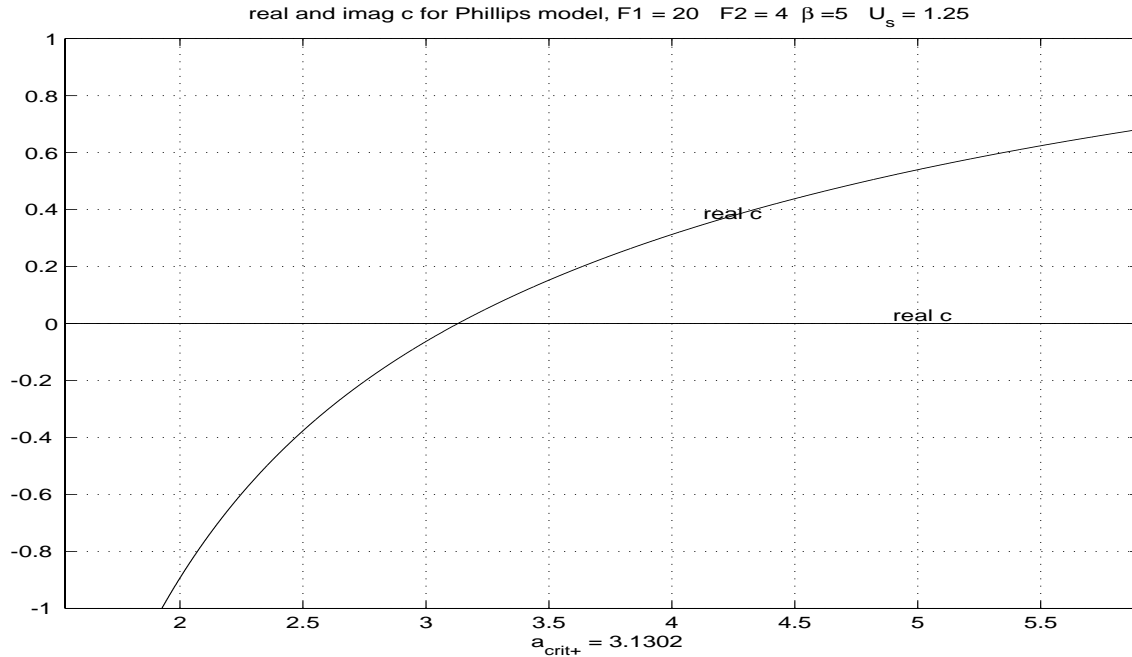
real and imag c for Phillips model, F1 = 20   F2 = 4 β =5 U_s = 1.25

a_{crit+} = 3.1302

Figure 10.4.4 c. The two roots for c as a function of a when the shear is positive and just equal to the minimum critical shear. Note the coalescence at c =0 the value of U_2.

d) The role of friction in the two-layer model.

The role of friction can be two-edged. On the one hand we expect friction to act as a dissipative mechanism and so imagine that it would be stabilizing. On the other hand in the inviscid problem with β the stabilization if provided by an inertial constraint (potential vorticity conservation) that might be weakened by friction. We shall see that friction can serve to both stabilize and destabilize the problem. To start simply, let's first consider the case in which the layer thicknesses are equal so that
$F_1 = F_2 = F$, $r_1 = r_2 = r$ and where we ignore the beta effect. Then from the inviscid dispersion relation (10.4.4) we would expect instability for all shears as long as the wavenumber $a < (2F)^{1/2}$. When friction is considered (but not β) the dispersion relation becomes,

$$c = U_2 + \frac{U_s}{2} - i\frac{r}{k}\frac{a^2 + F}{a^2 + 2F} \pm i\frac{\left[U_s^2 a^4 (4F^2 - a^4) + F^2 r^2 a^4 / k^2\right]^{1/2}}{2a^2(a^2 + 2F)} \qquad (10.4.12)$$

Note that if the radicand is positive, the real part of the phase speed is always

$$c_r = \frac{U_1 + U_2}{2} \qquad (10.4.13)$$

The condition that the imaginary part of c vanish can be obtained by equating the two final terms in (10.4.6). A little algebra leads to the critical condition,

$$U_s = \frac{ra}{k\sqrt{2F - a^2}} \qquad (10.4.14)$$

Note that the condition depends on k and l separately and not just on a. The marginal curve is plotted in figure 10.4.5a.
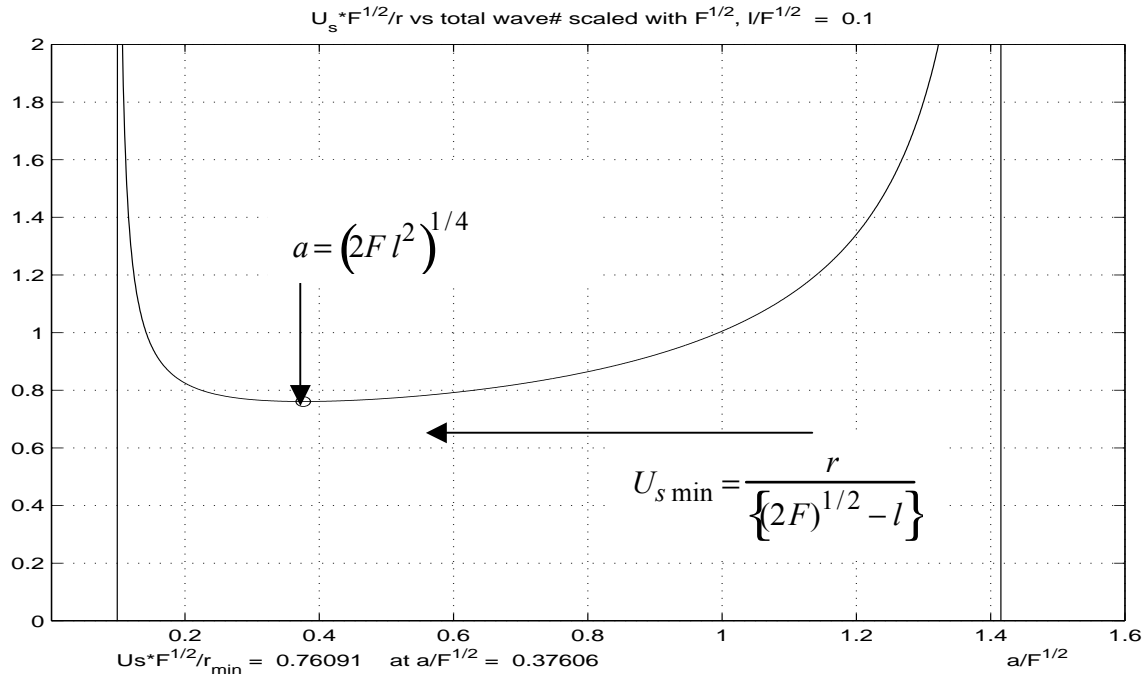
$a = \sqrt{2F}$

$a = l$

Figure 10.4.5a The neutral stability curve for the two layer model with friction but no beta.

The effect of friction introduces a minimum critical shear that must be exceeded before the energy release from the basic flow to the disturbance overbalances the energy of the perturbations lost to frictional dissipation. The imaginary part of c for large wavenumbers, where the flow is stable, is given by the first term in r on the right hand side of (10.4.6). When there is friction in the problem the coefficients of the differential equation are no longer real and so the complex solutions for c are no longer complex conjugates. Figure 10.4.5b shows the real and imaginary parts of the frequency as a function of wave number for a shear about 1/1/2 times the minimum critical value.
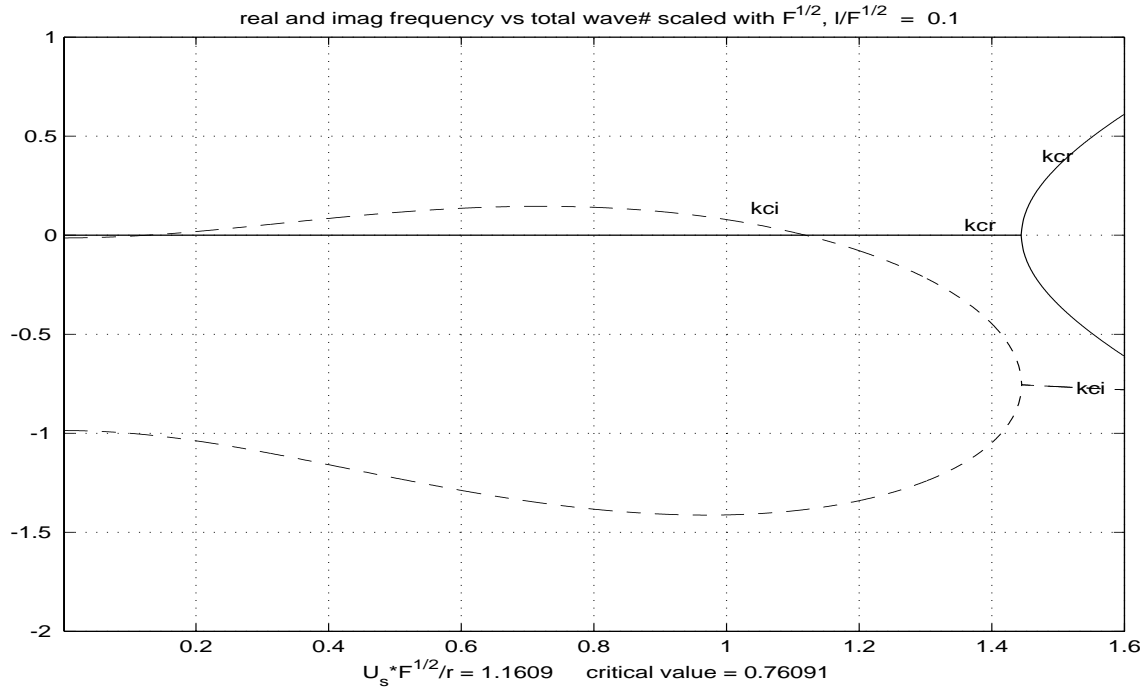
Figure 10.4.5 b The real and imaginary parts of the *frequency kc*.

The emergence of two roots for c_i occurs for negative values of the growth rate. As the wavenumber is decreased one root remains O(1) and negative while the other root goes through zero at the marginal curve. Note that in distinction to the inviscid case (e.g. figure 10.4.4b) the zero crossing of the growth rate is linear and does not have the square root behavior of the inviscid problem. For the viscous problem the marginal curve does not correspond to a coalescence of roots. Correspondingly, the region below the marginal curve in Figure 10.4.5a corresponds to *damped* modes, with negative values of growth rate. This is in distinction to the inviscid case where the stable modes were neutral, neither growing or decaying.

It is interesting to examine the energy balance in the perturbation. First note that from (10.3.8a) the amplitude ratio

$$\frac{A_2}{A_1} = \frac{a^2 + F}{F} + \frac{U_1 - U_2}{c - U_1} + ir\frac{a^2}{2k(c - U_1)F}$$

(10.4.15)

$$= 1 + \frac{a^2}{F} - \frac{U_s}{U_s/2 - ic_i} + ir\frac{a^2}{2k(U_s/2 - ic_i)F}$$

Note that the phase shift between the layers is due both to the growth rate, i.e. to c_i and to the dissipation r. Even if there is no growth there must be a phase shift to allow for the extraction of energy from the basic flow to balance dissipation. Let us examine that more carefully. At the marginal curve $c_i = 0$ and the shear is related to the friction coefficient by (10.4.14). A little algebra shows that on the marginal curve,

$$\left.\frac{A_2}{A_1}\right)_{c_i=0} = -1 + \frac{a^2}{F} - i\left(2F - a^2\right)^{1/2}\frac{a}{F}$$

(10.4.16)

In particular note that the magnitude of the amplitude ratio is *unity* on the marginal curve, that is,

$$\left|\frac{A_2}{A_1}\right|^2 = 1 - 2\frac{a^2}{F} + \frac{a^4}{F^2} + (2F - a^2)\frac{a^2}{F^2} = 1$$

(10.4.17)

The magnitude of the amplitudes, at the marginal curve, is the same in both layers and the amplitudes differ only by a phase shift. Thus,

$$A_2 = A_1 e^{i\varphi}$$

(10.4.18a)

where

$$\sin\varphi = -\left(2F - a^2\right)^{1/2} a/F$$

(10.4.18b)

so that the upper wave is shifted westward with height (as we know it must to release energy). To form the energy equation we multiply (10.2.5) by ϕ_n and integrate over the volume of the fluid to obtain,

$$\frac{\partial}{\partial t}\int_0^1\left[\sum_n\overline{|\nabla\phi_n|^2}/2 + F\overline{(\phi_1-\phi_2)^2}/2\right]dy = U_s\int_0^1\overline{F\phi_1\phi_{2x}}\,dy - \frac{r}{2}\int_0^1\sum_n\overline{|\nabla\phi_n|^2}\,dy \qquad (10.4.19)$$

The first term on the right hand side is the baroclinic energy conversion associated with the horizontal buoyancy flux. At the marginal curve the phase speed is $U_1 + U_2$ which we may take to be zero with no loss of generality. Then, since

$$\overline{\phi_1\phi_{2x}} = \frac{1}{2}\overline{\left(A_1e^{ikx} + A_1e^{-ikx}\right)\left(\frac{ik}{2}\right)\left(A_2e^{ikx} - A_2e^{-ikx}\right)}\sin^2 ly$$

$$= \frac{ik}{4}\left(A_1^*A_2 - A_2^*A_1\right)\sin^2 ly, \qquad (10.4.20)$$

$$= \frac{ik}{4}|A_1|^2\left\{e^{i\varphi} - e^{-i\varphi}\right\}\sin^2 ly = \frac{-k}{2}|A_1|^2\sin\varphi\sin^2 ly$$

the baroclinic energy conversion term is, using the identities valid on the critical curve,

$$F\frac{U_s}{2}k|A_1|^2\left\{(2F-a^2)^{1/2}\,a/F\right\}\int_0^1\sin^2 ly\,dy$$

$$\qquad (10.4.21)$$

$$= ra^2/4|A_1|^2$$

The last equality shows that the energy release by the baroclinic instability exactly balances the dissipation of energy since the last term on the right hand side of (10.4.19).

Above the marginal curve, where the flow is unstable there is an excess of energy released over that which is dissipated and that excess provides the energy for growth.

e) Instability with friction and β.

The situation becomes a little surprising when we add both friction and the beta effect. Again, to keep matters as algebraically as simple as possible we will take the case of equal layer thicknesses and no topography. The dispersion relation (10.3.10) now becomes

$$
c = \frac{U_1 + U_2}{2} - \frac{\beta(a^2 + F)}{a^2(a^2 + 2F)} - ir\frac{(a^2 + F)}{a^2(a^2 + F)}\frac{a^2}{F}
$$
$$
\pm \frac{\left[U_s^2 a^4 \left(a^4 - 4F^2 \right) + 4F^2 \left(\beta + ira^2/2k \right)^2 \right]^{1/2}}{2a^2(a^2 + 2F)}
\tag{10.4.21}
$$

The dispersion relation is rendered difficult because of the imaginary term within the radical. In particular,

$$
ic_i = -i\frac{(a^2 + F)}{a^2(a^2 + 2F)}\frac{ra^2}{2k} + i\,\mathrm{Im}\frac{\left[-U_s^2(4F^2 - a^4) + 4F^2\beta^2 + i8\beta F^2\, ra^2/2k \right]^{1/2}}{2a^2(a^2 + 2F)}
\tag{10.4.22}
$$

The radicand is of the form $A + iB = \mathrm{Re}^{i\vartheta}$ where

$$
A = -U_s^2 a^4 (4F^2 - a^4) + 4F^2\beta^2 - r^2 F^2\, a^4/k^2
$$
$$\tag{10.4.23 a, b}$$
$$
B = 8\beta F^2\, a^2/2k
$$

Since, $R = \left(A^2 + B^2 \right)^{1/2}$ and $\sqrt{A + iB} = R^{1/2}(\cos(\vartheta/2) + i\sin(\vartheta/2))$, it follows that

Chapter 10   31

$$\text{Im}(A + iB)^{1/2} = R^{1/2} \sin(\vartheta/2)$$

$$= \left(A^2 + B^2\right)^{1/4} \left\{ \frac{1 - \cos\vartheta}{2} \right\}^{1/2} \tag{10.4.24}$$

$$= \left\{ \frac{(A^2 + B^2)^{1/2} - A}{2} \right\}^{1/2}$$

With A and B defined by (10.4.23 a, b) it is straightforward to find the condition at the marginal curve for which $c_i = 0$. Letting $\gamma = ra^2/k$, a little algebra yields the condition,

$$4\gamma^4(a^2 + F)^4 + 4A\gamma^2(a^2 + F)^2 = B^2 \tag{10.4.25}$$

With substitutions for A, B and γ we obtain the equation for the marginal curve,

$$U_s^2 = \frac{4\beta^2 F^2}{a^2(a^2 + F)^2(2F - a^2)} + \frac{r^2 a^2}{k^2(2F - a^2)} \tag{10.4.26}$$

which should be compared to the equivalent curve for the completely inviscid case $r = 0$ obtained from (10.4.6). We denote that critical curve in the inviscid case by U_{sI},

$$U_{sI}^2 = \frac{4\beta^2 F^2}{a^4(4F^2 - a^4)} \tag{10.4.27}$$

If $\beta = 0$ we recover the stability curve given by (10.4.14). On the other hand as $r \to 0$ the limiting form for (10.4.26) is *not* (10.4.27), the inviscid curve. Indeed, the ratio,

$$\frac{U_s{}^2}{U_{sI}{}^2} = 1 - \frac{F^2}{(a+F)^2} + \frac{r^2 a^2}{k^2} \frac{a^4(2F+a^2)}{4\beta^2 F^2} \tag{10.4.28}$$

so that in the limit as $r \to 0$ $U_s / U_{sI} = \left(1 - a^2/F\right)^{1/2} < 1$! This implies that for a small amount of Ekman friction will destabilize the inviscid problem by destabilizing shears below the inviscid curve that were stable in the absence of friction. This seems a paradoxical result because we would like to believe that if the dissipation is small enough we ought to be able to ignore it and yet here we find that in the limit of vanishing friction the neutral curve is moved to lower shears by *an order 1 amount* , i.e. an amount independent of r as r goes to zero. Figure 10.4.6 shows the two marginal curves for a case of small r.
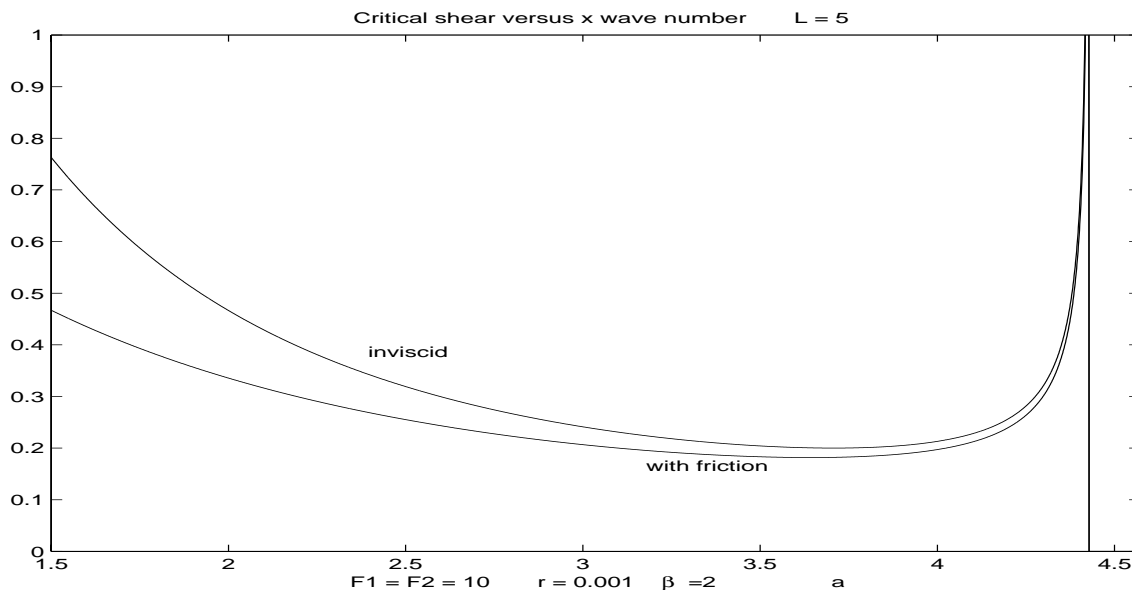

Figure 10.4.6 The inviscid and frictional marginal curves. For the frictional curve r =0.001.

It is clear that there is a region of shear between the two curves which has been destabilized by friction and this region, for small r is independent of the value of r. How

Chapter 10   33

is this possible if we believe in the physical continuity of the dynamics? The paradox is easily resolved if we examine the growth rate as a function of shear in the intervening region. Figure 10.4.7 shows the form that curve takes (a large value of $r = 0.01$)is used for the graph to make the plot clearer).
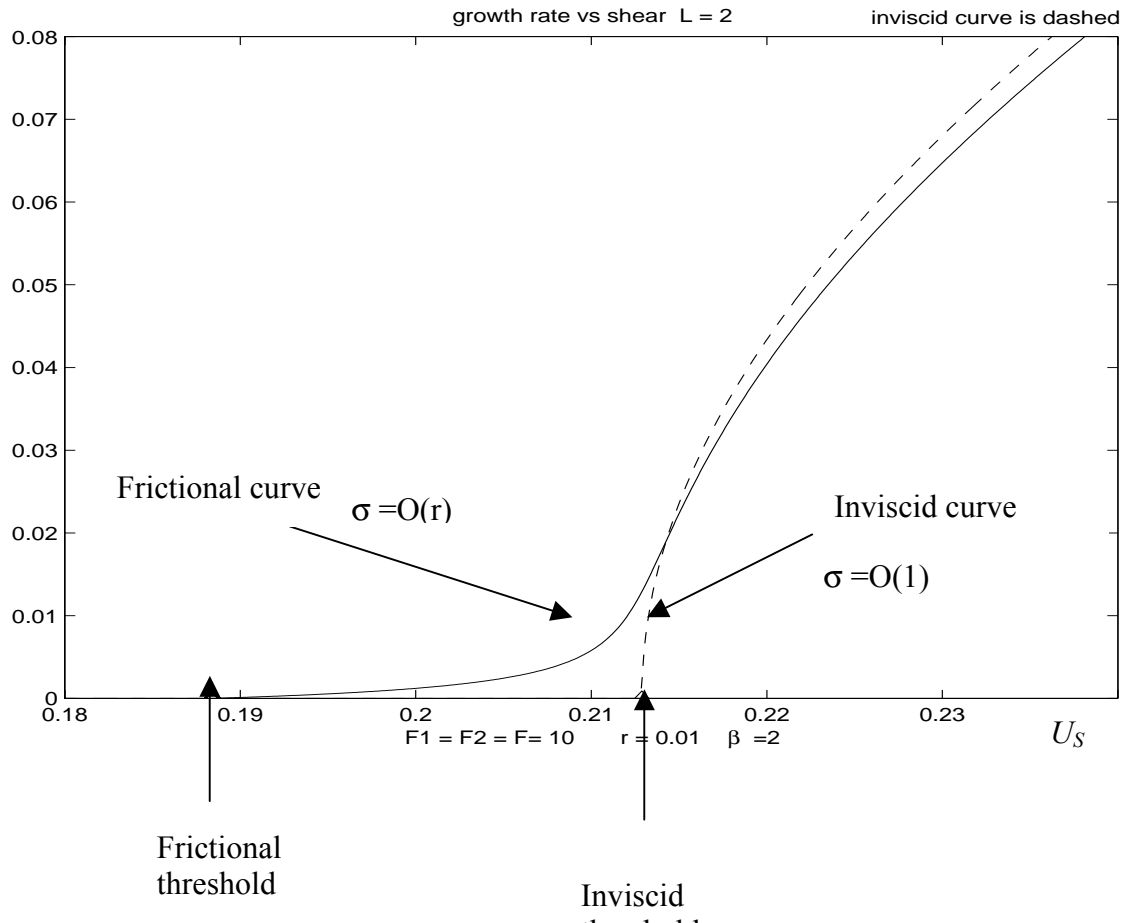

Figure 10.4.7 The growth rate as a function of shear, $r = .01$

Figure 10.4.7 shows the growth rate as a function of shear for a fixed wavenumber. As expected from Figure 10.4.6 the growth rate becomes positive at the threshold for the frictional curve first. In the intermediate region between the two thresholds the growth rate remains small $O(r)$ and only becomes $O(1)$ as the inviscid threshold is passed. The paradox is resolved by noting that the growth rate is $O(r)$ in the intervening region and hence the growth rate goes to zero as $r \rightarrow 0$ even while that region is formally a region of

weak instability. For larger shears the curves approach one another although note that the growth rate of the frictional case is always slightly smaller than the inviscid theory.

The physical cause of this rather strange behavior is due to the double nature of the role of friction. On the one hand friction, as we saw in the previous section, is responsible from providing a phase shift in the wave and that phase shift can liberate energy through that phase shifted structure of the disturbance. The friction also directly dissipates energy. When the friction is small, the amplitude variation provided by the β effect when phase shifted by the small friction releases more energy than is dissipated by friction. On the other hand, if the friction is large the dissipative effect is dominant and the friction is stabilizing as we would naively expect. Figure 10.4.8 shows the inviscid and frictional curves for a larger value of r.
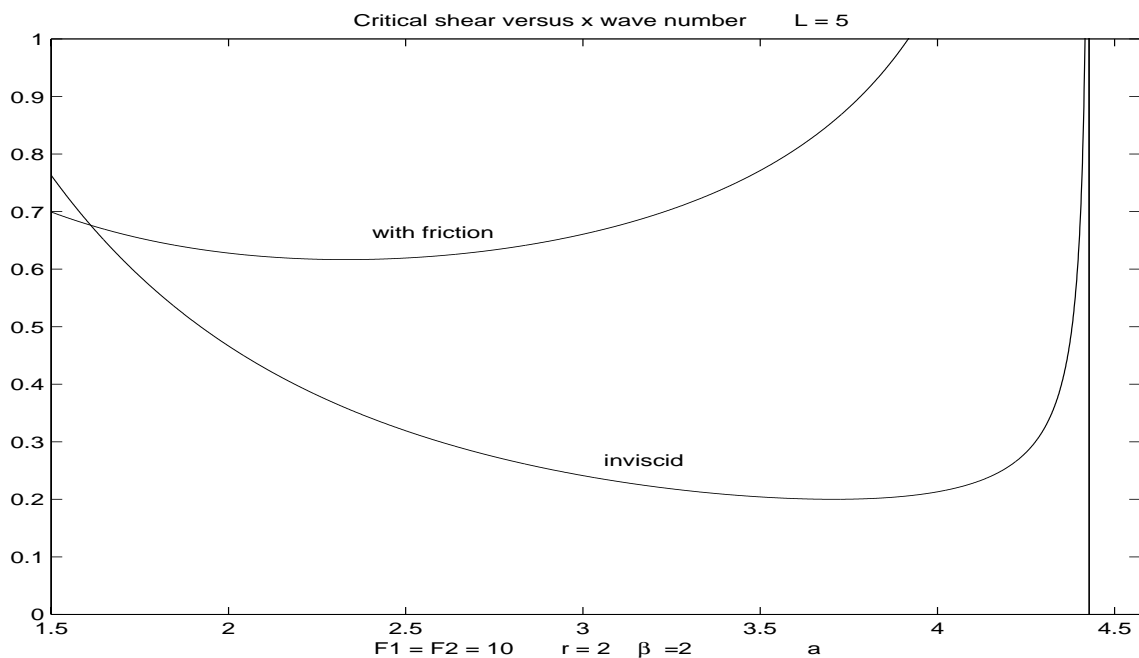

Critical shear versus x wave number   L = 5

with friction

inviscid

F1 = F2 = 10   r = 2   β =2   a

Figure 10.4.8 The neutral curves for the inviscid and frictional case when r =2.

In this case, with r =2 it is clear that friction is large enough to stabilize the flow. This situation frequently arises in stability theory. If a flow is rendered stable because of inertial constraints a small amount of dissipation can actually destabilize the flow although as in the baroclinic case the resulting growth rates are usually small. For the

baroclinic problem the result was discovered by Holopainen (1961, Tellus, **13**, 363-367) and extended to the finite amplitude problem by Romea (J. Atmos. Sci., **34**, 1689-1695). The classical example is the instability of pipe flow where the parabolic profile has a vorticity gradient that is always positive and hence, according to inviscid theory is always stable but is observed to be unstable for small enough but non zero friction. That problem, of great mathematical subtlety, is described by C.C. Lin in his monograph (Theory of Hydrodynamic Stability, Cambridge Univ. Press. 1955, pp155) and reviewed in Drazin and Reid (Hydrodynamic stability, Cambridge Univ. Press,1981, pp 527).