# Chapter 5: Normal mode Instability

## 5.1 Introduction

The linear partial differential equation (4.1.2) governing the instability of zonal flows has coefficients that depend on y and z but independent of x and t. To approach the solution of a general initial value problem where the perturbation field is specified at t=0 as a function of x, y, and z we could use a Laplace transform in time and a Fourier transform in x since the x interval is infinite (in our idealization) or cyclically periodic. We will return to the issue of the initial value problem later. However, it ought to be clear that the inversions, both Laplace and Fourier, that this approach implies would be very difficult to execute.

Instead a simpler approach is usually taken. In place of the initial value problem we abandon the initial conditions and search for the *normal modes* at fixed frequency, $\omega$ and x wavenumber k. i.e. a solution that is periodic in x (satisfying the finiteness boundary condition at plus and minus infinity in x. For simplicity in our formulation we will consider only the inviscid and adiabatic problem but it should be clear how to add dissipation to the discussion. So, since the coefficients of (4.1.2) are independent of t and x we ask whether solutions exist of the form

$$\phi = \Phi(y,z)e^{i(kx-\omega t)} \tag{5.1.1}$$

where the real part of the right hand side is implied.

Returning to (4.1.2) and inserting the above form, we first note that

$$\frac{\partial}{\partial t} \rightarrow -i\omega, \quad \frac{\partial}{\partial x} \rightarrow ik \tag{5.1.2}$$

It is helpful to define the phase speed *c*

$$c = \omega / k \tag{5.1.3}$$

The resulting partial differential equation for $\Phi$ is, with its boundary conditions,

$$(U_o - c)\left[\frac{1}{\rho_s}\frac{\partial}{\partial z}\left(\rho_s \frac{f_o^2}{N^2}\frac{\partial \Phi}{\partial z}\right) + \frac{\partial^2 \Phi}{\partial y^2} - k^2 \Phi\right] + \frac{\partial q_o}{\partial y}\Phi = 0,$$

$$(U_o - c)\Phi_z - U_{oz}\Phi + \frac{N^2}{f_o^2}(f_o h_y)\Phi = 0, \quad z = 0, z_t \qquad \text{(5.1.4 a,b,c)}$$

$$\Phi = 0, \qquad y = y_1, y_2$$

We can make the following observations:

1) The problem is homogeneous in the equations and the boundary conditions. Thus $\Phi$ =0 is always a possible solution. Indeed, for most values of $k$ and $\omega$ it will be the only solution. Special, non trivial solutions may exist for a given $k$ if $c$ takes on a particular value. For such solutions $c = c(k)$ . This is, therefore and eigenvalue problem for $c$ but it is not of the standard Sturm –Liouville form, $(p\Phi')' + (q + \lambda r)\Phi = 0$ where $p$ and $r$ are of a single sign in the interval of the problem. In fact, from a mathematical point of view the necessary conditions for instability are equivalent to the condition that (5.1.4) *not* be a standard S.L. problem. For standard S.L. problems the eigenvalue can always be shown to be real. We are interested in cases, where the eigenvalue is complex. The student should review the theorem that shows $\lambda$ is real for the S.L. problem (it requires $r$ to not change sign) and note how this is connected to our condition on the pv gradient for instability.
2) The interest in complex eigenvalues derives from the fact that if we insist that $k$ is real { we will examine later what happens if we relax this condition in a semi-infinite domain in x} but if $c$ is complex,

$$c = c_r + ic_i \tag{5.1.5}$$

then $e^{i(kx-kct)} = e^{ik(x-c_r t)}e^{kc_i t}$ . If the imaginary part of $c$, $c_i$ is positive we will have instability with *exponential growth*. The *growth rate* is $kc_i$

3) Since all the coefficients in the normal mode equation (5.1.4) are real it follows that if $(\Phi, c)$ is a solution to the eigenvalue problem, then $(\Phi^*, c^*)$ will also be a solution where * represents the complex conjugate of the signed variable. Therefore, if we find a solution that has $c_i < 0$ we are guaranteed that another solution, its complex conjugate exists with $c_i > 0$. For the inviscid problem then, instability is associated with complex values of the phase speed and exponential stability requires that all solutions have real $c$ .
4) It is not clear whether the spectrum, i.e. the set of normal modes, will be complete. That is, whether for each *k* there are enough modes in y and z to represent arbitrary initial data. Generally, the normal modes *do not* form a complete set and often for each *k* we might find one or two normal modes. That leaves the question open about the remaining solutions of the original problem, an issue we will return to.
5) Let's suppose that we have found , for each *k* a certain set of J normal modes, each one labeled $\{\Phi_j(y,z,k), c_j(k)\}, j = 1..J$ . Then the portion of the solution that can be carried by the normal modes must be synthesized for all *k* by the Fourier integral

$$\phi = \sum_{j=1}^{J} \int_k dk\, \Phi_j(y,z,k) e^{i(kx - c_j(k)t)} \tag{5.1.6}$$

**5.2 The relation to the necessary conditions for instability.**

In the last chapter we defined the Lagrangian displacement in the meridional direction by the relation,

$$v' = \phi_x = \left(\frac{\partial}{\partial t} + U_o \frac{\partial}{\partial x}\right)\eta \tag{5.2.1}$$

In accordance with our normal mode form, we write,

$$\eta = \operatorname{Re} Y(y,z) e^{ik(x-ct)} = \frac{1}{2}\left[ Y e^{i\theta} + Y^* e^{-i\theta^*} \right]$$

$$\theta \equiv k(x - ct) \tag{5.2.2}$$

where $\theta$ is the wave phase (complex). From (5.2.2) and (5.1.1) it follows that,

$$\Phi = (U_o - c)Y \tag{5.2.3}$$

In the stability theorems of Chapter 4 a key role was played by the x-averaged dispersion $\overline{\eta^2}$. Let's calculate that using (5.2.2) . First note that

$$\overline{e^{i2\theta}} = 0, \qquad \theta - \theta^* = -k(c - c^*)t = -i2kc_i t \tag{5.2.4}$$

Thus,

$$\begin{aligned} \overline{\eta^2} &= \frac{1}{4} \overline{\left[ Y e^{i\theta} + Y^* e^{-i\theta^*} \right] \left[ Y e^{i\theta} + Y^* e^{-i\theta^*} \right]} \\ &= \frac{2}{4} |Y|^2 e^{2kc_i t} \\ &= \frac{1}{2} \frac{|\Phi|^2}{|U_o - c|^2} e^{2kc_i t} \end{aligned} \tag{5.2.5}$$

We have, therefore, the simple relation,

$$\frac{\partial}{\partial t} \overline{\eta^2} = kc_i \frac{|\Phi|^2}{|U_o - c|^2} e^{2kc_i t} > 0 \tag{5.2.6}$$

This illustrates directly (for the normal mode solutions) that the Lagrangian displacements must increase with time if the solution is unstable. It also allows us to directly translate our conditions for instability to the normal modes, e.g. (4.2.23) becomes,

$$kc_i\left[\iint dydz\rho_s \frac{q_{o_y}}{2}\frac{|\Phi|^2}{|U_o-c|^2}+\int dy\rho_s \frac{|\Phi|^2}{|U_o-c|^2}\left(\tfrac{f_o^{\,2}}{N^2}U_{o_z}-f_oh_y\right)/2\Bigg\}_{z=0}^{z=z_t}\right]=0 \qquad (5.2.7)$$

Thus, if the growth rate is to be different from zero the integrals within the square brackets must add to zero and this yields the same conditions on the basic state for instability as derived in Chapter 4. In this form (5.2.7) is often called the *Charney-Stern* theorem although the original derivation lacked the boundary terms.

It is important to note that the equation (5.1.4) is singular on line in the y-z plane where $U_O - c = 0$. If $c$ is real this may occur for real values of y and z. That line is called the *critical line*. For problems in which either z or y can be removed from the problem (when $U_O$ is a function of only z or y) the position of the singularity is called the *singular point* of the equation. The presence of the singularity renders the problem very delicate especially in the interesting situation in which $c$ is just real, i.e. on the parametric boundary of instability. We are often particularly interested in determining the critical values of the parameters like β or $k$ that just render the flow unstable, i.e. on the border between stability and instability, on the border between real and complex $c$ . These are the curves in some parametric space e.g. $\beta = \beta(k)$ called the curve of marginal instability. We have to expect, generally, a rather delicate mathematical problem if such values yield critical lines or points to the differential equation. This in turn raises the important question as to whether the phase speed lies within the range of values of $U_O$ so that $U_O - c = 0$ will occur for some line in the y-z plane. What can we say *a priori* about the values of $c$ to be expected from the eigenvalue problem, especially for unstable modes.

**5.3 Bounds on the phase speed.**

We can gather some advanced information about the phase speed by rewriting the stability equation (5.1.4) in terms of the Lagrangian displacement amplitude *Y*. With (5.2.3) we find,

$$\Phi_y = (U_o - c)Y_y + U_{oy}Y,$$

$$\Phi_{yy} = (U_o - c)Y_{yy} + 2U_{oy}Y_y + U_{o_{yy}}Y = \frac{\frac{\partial}{\partial y}\left[(U_o - c)^2 Y_y\right]}{(U_o - c)} + U_{oyy}Y$$

(5.3.1)

Similarly,

$$\left(\rho_s \frac{f_o^2}{N^2}\Phi_z\right)_z = \frac{\left[(U_o - c)^2 \rho_s (f_o^2/N^2) Y_z\right]_z}{(U_o - c)} + Y\left(\rho_s \frac{f_o^2}{N^2} U_{oz}\right)_z \qquad (5.3.2)$$

Note too that,

$$(U_o - c)\Phi_z - U_{oz}\Phi = (U_o - c)^2 Y_z \qquad (5.3.3)$$

If we now rewrite the eigenvalue problem (5.1.4) in terms of *Y* we obtain,

$$\frac{\partial}{\partial y}\left(\rho_s [U_o - c]^2 Y_y\right) + \frac{\partial}{\partial z}\left(\rho_s \frac{f_o^2}{N^2}[U_o - c]^2 Y_z\right) - k^2 \rho_s [U_o - c]^2 Y + \rho_s \beta [U_o - c]\ Y = 0$$

(5.3.4)

with the boundary conditions,

$$[U_o - c]\ Y_z + \frac{N^2}{f_o} h_y Y = 0, \qquad z = 0, z_t$$

$$Y = 0, \qquad y = y_1, y_2 \qquad (5.3.5\text{ a,b})$$

Now multiply (5.3.4), integrate over the y,z domain of the problem and use the boundary conditions (5.3.5 a,b) to obtain:

$$\int_0^{z_t}\int_{y1}^{y2} dy\,dz\left\{(U_o-c)^2P-\beta(U_o-c)J\right\}+\int_{y1}^{y2}(U_o-c)f_oh_yJdy\Bigg|_{z=0}^{z=z_t}=0,$$

(5.3.6 a,b,c)

$$P=\rho_s\left[\left|Y_y\right|^2+\frac{f_o{}^2}{N^2}\left|Y_z\right|^2+k^2\left|Y\right|^2\right],\quad J=\rho_s\left|Y\right|^2$$

Note that *P* and *J* are each positive definite. Now let's consider the real and imaginary parts of (5.3.6a) separately. The imaginary part is,

$$c_i\left\{\int\int dydz\left[(U_o-c_r)P-\frac{\beta}{2}J\right]+\int dy\frac{J}{2}f_oh_y\Big|_0^{z_t}\right\}=0, \tag{5.3.7a}$$

while the real part yields,

$$\int\int dydz\left[\left((U_o-c_r)^2-c_i{}^2\right)P-\beta(U_o-c_r)J\right]+\int dyJ(U_o-c_r)f_oh_y\Big|_0^{z_t}=0$$

(5.3.7b)

*Y* is a continuous function that vanishes at the points *y1* and *y2* and so it can be represented by a Fourier sine series that can be differentiated term by term (because of the continuity and the appropriate boundary conditions).

Thus,

$$Y=\sum_{j=1}Y_j\sin(j\pi[y-y_1]/L),\quad L=y_2-y_1, \tag{5.3.8a}$$

and

$$Y_y = \sum_{j=1} Y_j (j\pi / L) \cos(j\pi[y - y_1]/L) \qquad (5.3.8.b)$$

Thus

$$\int_{y_1}^{y_2} |Y|^2 dy = \frac{L}{2} \sum_{j=1} |Y_j|^2,$$

$$\int_{y_1}^{y_2} |Y_y|^2 dy = \frac{L}{2} \sum_{j=1} |Y_j|^2 \frac{j^2 \pi^2}{L^2} \leq \frac{L}{2} \frac{\pi^2}{L^2} \sum_{j=1} |Y_j|^2 = \frac{\pi^2}{L^2} \int_{y_1}^{y_2} |Y|^2 dy \qquad (5.3.9)$$

It follows therefore that,

$$\overline{P}^{y,z} \geq (k^2 + \pi^2 / L^2) \overline{J}^{y,z} \qquad (5.3.10)$$

where the overbar y,z label denotes integration over the y,z domain.

From (5.3.7a) *if $c_i$ is not zero, i.e. for an unstable wave*,

$$c_r = \frac{\overline{U_o P}^{y,z}}{\overline{P}^{y,z}} - \frac{\beta}{2} \frac{\overline{J}^{y,z}}{\overline{P}^{y,z}} + \frac{1}{2} \frac{\left. \int_{y_1}^{y_2} J f_o h_y dy \right|_0^{z_t}}{\overline{P}^{y,z}} \qquad (5.3.11)$$

We suppose that the basic flow smoothly varies between its minimum and maximum values over the (y,z) domain. Let's restrict attention to the case where there is only a sloping surface at the fluid's bottom, i.e. at z=0. Then using (5.3.10) and (5.3.11) it follows directly that,

$$c_r > U_{o\min} - \frac{\beta}{2\left[k^2 + \pi^2/L^2\right]} - \frac{f_o h_{y\max}}{2} \frac{\left. \int_{y1}^{y2} J\, dy \right|_{z=0}}{\bar{J}^{y,z}\left[k^2 + \pi^2/L^2\right]} \tag{5.3.12}$$

Unfortunately, the last term involves two integrals over the function *Y* and so it depends on the vertical structure of the mode. Let's suppose that we can write

$$\bar{J}^{y,z} = L_z \int_{y1}^{y2} J(y,0)\, dy \tag{5.3.13}$$

and so define a vertical scale for *Y*. Then,

$$c_r > U_{o\min} - \frac{\beta}{2\left[k^2 + \pi^2/L^2\right]} - \frac{f_o h_{y\max}/L_z}{2\left[k^2 + \pi^2/L^2\right]} \tag{5.3.14 a}$$

Similarly, we can easily show that, if $h_y > 0$,

$$c_r \le U_{o\max} \tag{5.3.14 b}$$

Thus, if we could ignore the β effect and if the bottom were flat, it would follow that the real part of the phase speed of all unstable waves would have to *lie within the range of the basic velocity* $U_O$. That guarantees that the stability equation, aside from certain exceptional cases to be discussed, must have a singularity in the real domain if *c* is real and on the edge of being unstable. Note the physical plausibility of this limit. If the phase speed could be much larger than the basic flow one might be able to argue that to lowest order the propagating wave would not be aware at all of the basic flow. As a first approximation one would ignore the flow altogether. In 12.802 this was a frequent linearization assumption. In that case it is impossible to see where an instability could come from if there is no basic flow in the problem. Clearly, the bounds developed above are tighter than that extreme argument but the physical basis for understanding the existence of the bounds is connected to the need of the growing wave to sense the velocity variation over its spatial structure.

To continue, let's for simplicity ignore the possibility of a bottom slope although it can be included if we are willing to use the device introduced in (5.3.14 a)

Returning to (5.3.11) we have,

$$\overline{U_o P}^{y,z} = c_r \overline{P}^{y,z} + \frac{\beta}{2}\overline{J}^{y,z}, \; c_i \neq 0 \tag{5.3.15 a}$$

while from (5.3.7b),

$$\overline{U_o^2 P}^{y,z} - 2c_r \overline{U_o P}^{y,z} + (c_r^2 - c_i^2)\overline{P}^{y,z} - \beta \overline{(U_o - c_r)J}^{y,z} = 0 \tag{5.3.15 b}$$

while using (5.3.15 a) this can be rewritten,

$$\overline{U_o^2 P}^{y,z} - (c_r^2 + c_i^2)\overline{P}^{y,z} = \beta \overline{U_o J}^{y,z} \tag{5.3.15 c}$$

Consider any two constant $a$ and $b$ and then construct the product,

$$\overline{(U_o - a)(U_o - b)P}^{y,z} = \overline{U_o^2 P}^{y,z} - (a+b)\overline{U_o P}^{y,z} + ab\overline{P}^{y,z} \tag{5.3.16}$$

Using the relations (5.3.15 a) and (5.3.15 c) we can write the above equation as,

$$\overline{(U_o - a)(U_o - b)P}^{y,z}$$

$$= (c_r^2 + c_i^2)\overline{P}^{y,z} - c_r(a+b)\overline{P}^{y,z} + ab\overline{P}^{y,z} - \frac{(a+b)}{2}\beta \overline{J}^{y.z} + \beta \overline{U_o J}^{y,z} \tag{5.3.17}$$

$$= \left\{[c_r - (a+b)/2]^2 + c_i^2\right\}\overline{P}^{y,z} - \left\{\frac{a-b}{2}\right\}^2 \overline{P}^{y,z} + \beta \overline{\{U_o - (a+b)/2\}J}^{y,z}$$

Equivalently, a slight rearrangement yields,

$$\left(c_r-(a+b)/2\right)^2+c_i^2=\left(\frac{a-b}{2}\right)^2+\beta\frac{\overline{\left[\frac{a+b}{2}-U_o\right]J}^{y,z}}{\overline{P}^{y,z}}+\overline{(U_o-a)(U_o-b)P}^{y,z}\Big/\overline{P}^{y,z}$$

(5.3.18)

and note that this is true for any two constants $a$ and $b$. By definition, in the y,z interval of our problem, $U_{\min}\le U_o\le U_{\max}$. Let's choose $a=U_{\max}$, $b=U_{\min}$. It therefore follows that the last term in (5.3.18) is always less than zero. We have then, using that fact and (5.3.10)

$$\left(c_r-[U_{\max}+U_{\min}]/2\right)^2+c_i^2\le\left[\frac{U_{\max}-U_{\min}}{2}\right]^2+\frac{\beta}{2}\frac{[U_{\max}-U_{\min}]}{k^2+\pi^2/L^2} \qquad (5.3.19)$$

This beautiful theorem in originally due to Howard (JFM 1961, vol.10 509-512)who proved it for the case of a nonrotating flow (β =0) in the context of the stability of stratified shear flows. Note that if β =0 the theorem says the complex phase speed must lie within a semicircle centered on the mean velocity and whose radius is half the maximum difference between the two velocities.

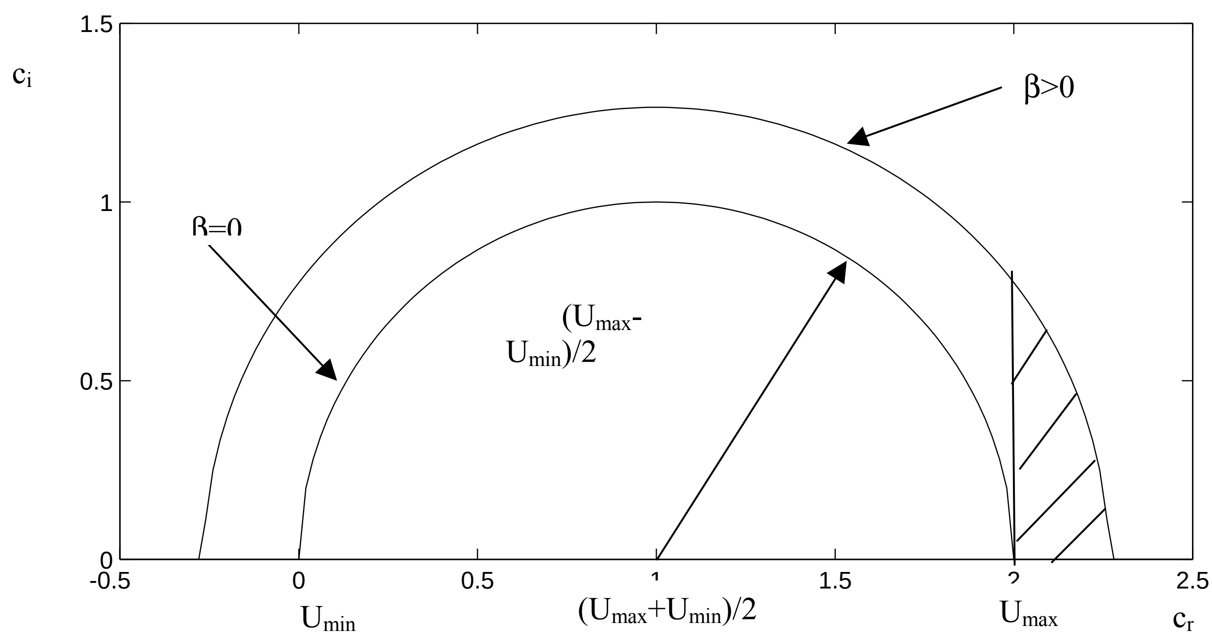

Figure 5.3.1 The semi-circle in which the complex phase speed must lie. The inner semi-circle obtains when β=0 and the outer one for β>0. The region cross-hatched is not allowed since the real part of the phase speed must be less than the maximum velocity of the basic flow.

Note that as the shear goes to zero the radius of the semicircle, even with β, shrinks to zero. There is no instability in the absence of shear.

The particular bound given by the semicircle theorem with β has the unpleasant feature of apparently allowing increased growth rates as β increases and this is counter-intuitive. In fact, the radius increases merely to accommodate the possibility of complex $c$ which lie in the small region less that $U_{\min}$.

Other bounds on the imaginary part of $c$ can be derived by similar methods. For example, if instead of the transformation (5.2.3) we introduce the function $\chi$ by,

$$\Phi = (U_o - c)^{1/2} \chi \qquad (5.3.20)$$

and derive the equation for $\chi$ and then exploit the same integral methods as we used in the semicircle theorem, it is not difficult to obtain the following bound on the growth rate:

$$(2kc_i)^2 \le \left[ U_{o_y}{}^2 + \frac{f_o{}^2}{N^2} U_{o_z}{}^2 \right]_{\max} \frac{k^2}{k^2 + \pi^2 / L^2} \qquad (5.3.21)$$

so that the growth rate is bounded by the maximum over the meridional domain of the flow of the available kinetic energy of basic flow, represented by the horizontal velocity shear and available *potential* energy represented by the second term in the square bracket $(b_{o_y})^2 / N^2$ where $b_{o_y}$ is the meridional gradient of the buoyancy in the basic flow. For details see Chapter 7 of GFD. Note too, that (5.3.21) implies that the growth rate goes to zero as $k \rightarrow 0$ for all finite $L$. For $k$=0, there will be no meridional geostrophic velocity in the perturbation and without the fluid parcels crossing the current they experience no change in the velocity field and so can't release either the kinetic energy stored in the horizontal shear or the available potential energy manifested by the horizontal buoyancy gradient. This means that the maximum in the growth rate will generally occur for scales that are O(1) compared to the intrinsic scales of the problem. They are i) either the scale of the shear of the current, or ii) an intrinsic scale like the deformation radius. We could draw that same conclusion from the semicircle theorem which implies that the imaginary part of $c$ is always finite and hence the growth rate, $kc_i$ must go to zero as $k$ vanishes.

In the next chapter we will take a specific example to illustrate these points.