DYNAMICAL INSTABILITIES IN STARS

P.LEDOUX

Institut d'Astrophysique

Université de Liège

SUMMARY

The linear dynamical instability at the origin of convection in stars is reviewed and shown to depend essentially on the sign of

$$A = \frac{1}{\rho}\frac{d\rho}{dr} - \frac{1}{\Gamma_1 p}\frac{dp}{dr}$$

which is the usual argument of convection criteria. The case of two or more superadiabatic regions separated by subadiabatic ones might well deserve more detailed attention.

Once this instability is partially removed by the setting in of convection its effects must be balanced by dissipation terms if a stationary state is to result. This yields the value of a Rayleigh number.

If energy generation is included in the non-conservative terms, possibilities are somewhat enriched including a case of dynamical instability in presence of A<0 (usually stable) but very small in absolute value.

1. THE GENERAL PROBLEM

In the context of this conference we are not interested in dynamical instability towards purely radial modes since convection cannot manifest itself through these modes. We are thus left with the problem of the response of the star to non radial perturbations which

https://doi.org/10.1017/S0252921100112333 Published online by Cambridge University Press

we shall assume very small to allow a linear treatment.

Of course, non linear effects may be of very great interest and importance and are at least partially included in some of the approaches to steady convection as, for instance, in the mixing length theory or various numerical attempts usually with other simplifying hypothesis and simple geometry (cf. Spiegel 1971, 1972 Nordlund, 1976). One of the most recent and most direct attacks in general stellar circumstances is due to Deupree (1974, 1975 a-b, 1976). As far as I am aware, it has not revealed new instabilities such as may occur for instance in metastable situations. It has not either restricted the domain of significant dynamical instabilities but it has yielded interesting information on the development of these instabilities such for instance as a strong asymmetry between upward and downward motions.

The study of non radial stellar oscillations does not go far back. If we exclude Lord Kelvin's discussion of the homogeneous incompressible sphere (Thomson, 1863) and some more or less timid references by Moulton (1909) and Shapley (1914), the first significant paper is that of Pekeris (1938) in which he solved the problem of the non radial perturbation of the homogeneous compressible sphere.

Pekeris used the usual separation in time and spherical coordinates of the Euclidian perturbations $f'(p',\rho', T',\Phi')$ and of the radial component of the displacement δr

$$f'(r,\theta,\phi,t) = f'(r)P_\ell^m(\cos\theta)e^{im\phi}e^{i\sigma t}, \quad -\ell\leqslant m\leqslant\ell$$

He showed that, for each value of the degree ℓ of the spherical harmonic, apart from a positive spectrum with an accumulation point at infinity corresponding to the pressure modes (or p modes, σ_p^2), there

https://doi.org/10.1017/S0252921100112333 Published online by Cambridge University Press

existed a negative spectrum corresponding to gravity modes (or g modes, σ_g^2) with an accumulation point at zero, all these modes being $(2\ell+1)$ degenerate with respect to m. An illustration of the distribution of the σ^2 as function of ℓ and the order of the modes can be found for instance in Ledoux (1974).

Of course, in this case, all the g modes ($\sigma_g^2<0$) are unstable but I don't think that the connection with convection was pointed out. Note that Pekeris choice of dependent variable ($\alpha=$ div $\delta\vec{r}$) led him to miss the so called fundamental mode (or f mode or Kelvin mode) which, in this case, is exactly the same as the unique mode (for a given ℓ) of the incompressible sphere ($\alpha=0$).

The next important paper is that of Cowling (1941) in which he tackled the case of the general polytrope of index $n(p= K\rho^{n+1/n})$. In this case, the general problem is of the fourth order as it does not split into two second order differential equations which can be solved successively as for the homogeneous model. However Cowling noted that, except for the lowest modes and lowest values of ℓ, the perturbation of the gravitational field can be neglected without serious effects. With this approximation, the problem reduces again to the second order and, as Cowling showed, can be assimilated to a Sturm-Liouville problem for large enough σ^2 (high p modes) or small enough $|\sigma^2|$ (high g modes), the f mode for each ℓ falling in between the corresponding p and g spectra.

Furthermore if the generalized ratio of specific heats Γ_1 (or γ in a pure gas) satisfies the inequalities

$$\Gamma_1 > \frac{n+1}{n} \tag{1}$$

all the σ_g^2 are positive (stable g modes : g^+) while, if

$$\Gamma_1 < \frac{n+1}{n} \tag{2}$$

they are all negative leading to instability (unstable g modes : g^-). But the criterion for convective instability in terms of

$$A = \frac{1}{\rho}\frac{d\rho}{dr} - \frac{1}{\Gamma_1 p}\frac{dp}{dr} \tag{3}$$

can be written very generally, even in the relativistic case (Thorne, 1966, Kovetz, 1967, Islam, 1970)

$$A > 0 \tag{4}$$

and becomes in a polytrope, with the usual notation,

$$A = \frac{1}{\theta}\frac{d\theta}{d\xi}\left(n - \frac{n+1}{\Gamma_1}\right) > 0 \tag{5}$$

or since $(d\theta/d\xi)$ is negative

$$\Gamma_1 < \frac{n+1}{n} \tag{6}$$

which is identical to condition (2). Thus all the σ_g^2 are negative (dynamical instability towards non radial perturbations) provided the criterion for convective instability (4) or (5) be satisfied everywhere in the star.

The negative g spectrum in the homogeneous model may be interpreted in the same way since in that case

$$A = -\frac{1}{\Gamma_1 p}\frac{dp}{dr} > 0$$

But apparently, it was only slowly (Ledoux, 1949), at least in astrophysics, that the connection between dynamically unstable g modes and the sign of A became to be recognized and that convection became to be considered as the end effect of this instability.

In the general case, it is not difficult to have A appear explicitly in the equations which can be written

$$\frac{\delta\rho}{\rho} = \frac{\rho'}{\rho} + \frac{\delta r}{\rho}\frac{d\rho}{dr} = -\operatorname{div}\vec{\delta r} \tag{7}$$

$$\sigma^2\vec{\delta r} - \operatorname{grad}\left(\phi' + \frac{p'}{\rho}\right) + \frac{\vec{r}}{r} A \frac{\Gamma_1 p}{\rho}\operatorname{div}\vec{\delta r} =$$

$$= \frac{\Gamma_3 - 1}{i\sigma}\frac{1}{\rho}\operatorname{grad}\rho\left(\varepsilon - \frac{1}{\rho}\operatorname{div}\vec{F}\right)' - \frac{i\sigma}{\rho}\operatorname{div}\mathscr{P}(\vec{\delta r}) \tag{8}$$

$$p' + \Gamma_1 p\left(\frac{\rho'}{\rho} + A\delta r\right) = \frac{\Gamma_3 - 1}{i\sigma}\rho\left(\varepsilon - \frac{1}{\rho}\operatorname{div}\vec{F}\right)' \tag{9}$$

$$T' - T\left(\frac{\Gamma_2 - 1}{\Gamma_2}\frac{p'}{p} + S\delta r\right) = \frac{1}{i\sigma C_p}\left(\varepsilon - \frac{1}{\rho}\operatorname{div}\vec{F}\right)' \tag{10}$$

$$\Delta\phi' = 4\pi G\rho' \tag{11}$$

where

$$\varepsilon = \varepsilon_N - \varepsilon_\nu + \frac{1}{\rho}\sum_{ik} P^{ik}\Delta_i v_k$$

represents the total heat liberated per second by nuclear reactions and viscosity.

One may note that A is related to the Brunt-Väisälä frequency N by

$$N^2 = -gA$$

https://doi.org/10.1017/S0252921100112333 Published online by Cambridge University Press

and to

$$S = \frac{\Gamma_2 - 1}{\Gamma_2} \frac{1}{p} \frac{dp}{dr} - \frac{1}{T} \frac{dT}{dr} \qquad (12)$$

by

$$A = \frac{4 - 3\beta}{\beta} S + \frac{1}{\mu} \frac{d\mu}{dr} \qquad (13)$$

if β is the ratio of the gas pressure to the total pressure.

The adiabatic approximation (right hand members neglected in (8), (9), (10)) is sufficient to discuss dynamical stability and it should enable us to understand the correlation noted above between the sign of A and the stability of the g modes.

If we neglect ϕ', we can for instance write a second order differential equation for $\xi = \delta r/r$:

$$\frac{d^2\xi}{dr^2} + \frac{d\xi}{dr} \left\{\frac{6}{r} + \frac{1}{\rho} \frac{d\rho}{dr} + \frac{1}{(\;)} \frac{d}{dr} \left(\frac{\rho r^2}{\Gamma_1 p}\right)\right\}$$

$$+ \xi \left\{ \frac{\sigma^2 \rho}{\Gamma_1 p} - \frac{\ell(\ell+1)}{r^2} \frac{\sigma^2 + Ag}{\sigma^2} + \frac{6}{r^2} - \frac{d}{dr} \left(\frac{g\rho}{\Gamma_1 p}\right) \right.$$

$$\left. + \frac{3}{r} \left\{\frac{1}{\rho} \frac{d\rho}{dr} + \frac{1}{(\;)} \frac{d}{dr} \left(\frac{\rho r^2}{\Gamma_1 p}\right)\right\} - \frac{1}{(\;)} \frac{g\rho}{\Gamma_1 p} \frac{d}{dr} \left(\frac{\rho r^2}{\Gamma_1 p}\right)\right\} = 0 \qquad (14)$$

where

$$(\;) = \left(\frac{\ell(\ell+1)}{\sigma^2} - \frac{\rho r^2}{\Gamma_1 p} \right)$$

For high p modes, σ^2 large, the equation simplifies very much and shows the acoustic character of these modes. However, they are without interest in the present context since, in realistic stel-

https://doi.org/10.1017/S0252921100112333 Published online by Cambridge University Press

lar conditions, all the σ_p^2 are always positive. But in the same way if one considers high g modes, $|\sigma^2|$ small so that terms proportional to σ^2 can be neglected, the only term left over which contains σ^2 is the second one in the coefficient of ξ in (14), i.e.

$$-\frac{\ell(\ell+1)}{r^2}\frac{Ag}{\sigma^2}$$

and the problem is essentially of the Sturm-Liouville type with a parameter $\lambda = 1/\sigma^2$. It is well known in that case that if A keeps the same sign throughout, the eigenvalue $(\lambda=1/\sigma^2)$ will be of the opposite sign. In other words, if A is positive (convective instability), at least the small σ_g^2 will be negative. But it has been accepted generally that all the σ_g^2 will have this same sign, because it is difficult to see how the sign of σ^2 could change in going from small to larger values if the sign of its coefficient is constant. Anyway fairly recently Grisvard, Souffrin and Zerner (1972) using the second order system

$$\frac{dv}{dr} = \left(\frac{\ell(\ell+1)}{\sigma^2} - \frac{\rho r^2}{\Gamma_1 p}\right)\frac{p^{2/\Gamma_1}}{\rho}\, w = aw \qquad (15)$$

$$\frac{dw}{dr} = \left(\sigma^2 + Ag\right)\frac{\rho}{r^2 p^{2/\Gamma_1}}\, v = bv \qquad (16)$$

equivalent to (14) managed to prove without any assumption as to the order of magnitude of σ^2 that the latter are all real positive if A is everywhere negative. Thus a necessary condition for dynamical instability is that A be positive at least in some part of the star. In that case, the authors succeeded in establishing an upper limit for the modulus of any negative σ_g^2

$$|\sigma_g^2| < \text{Max (Ag)} \tag{17}$$

where the maximum is taken on the region where A>0. They showed also that $|\sigma_g^2|$ increases with ℓ, i.e. when the horizontal wave length decreases.

However if the investigation of Grisvard et al is free of any asymptotic restriction, it still treats the perturbation of the gravitational potential ϕ' as negligible. Would the previous results hold if ϕ' is not neglected, especially for the lowest modes and harmonics? Lebovitz (1965 a-b, 1966) tackled this general problem on the basis of the variational principle established earlier by Chandrasekhar (1964), which expresses σ^2 as the extremum of integrals extended to the whole configuration. This enabled him to show first that if A<0 everywhere (convective stability) all the σ^2 are positive which is thus a sufficient condition for dynamical stability towards non-radial perturbations. It is also a necessary condition as he showed later, since the existence of a region however small with A>0 entails negative eigenvalues. This implies also that one can find solutions of the differential equations with appreciable amplitudes in that region only.

In this respect, the work of Ledoux and Smeyers (1966) was more or less complementary since they pointed out indeed that when A changes sign in $0 \leq r \leq R$ the asymptotic form (σ^2 small, g modes) of equation (14) has a turning point in A = 0 and that the g spectrum splits then into two : one of positive eigenvalues ($\sigma_g^2 > 0$) corresponding to modes oscillating in space with appreciable amplitudes in the convectively stable region (A<0) and decaying exponentially in the unstable region (A>0) while the other spectrum of negative eigenvalues ($\sigma_g^2 < 0$)

https://doi.org/10.1017/S0252921100112333 Published online by Cambridge University Press

corresponds to modes oscillating in the convectively unstable region (A>O) and decaying in the stable one.

Using equations (15) and (16), Scuflaire (1974, cf. also Osaki 1975) showed that these properties of the unstable modes subsist even for the first few modes, i.e. they oscillate only in the unstable region. However, it may happen that stable modes oscillate in a slightly superadiabatic region, as for instance in a condensed model where the stable g modes characteristic of the central stable core may continue to oscillate in an external convection zone.

2. MULTIPLE UNSTABLE ZONES

The unstable modes are of greatest interest here and an interesting problem arises as to their behaviour if there are two or more regions with A>O separated by stable zones (A<O). Tassoul and Tassoul (1968) in a discussion of asymptotic g modes suggested that there should be as many distinct unstable g spectra as there are unstable regions.

However, this is not obvious since even in the simplest case of two turning points (for instance two convectively unstable regions separated by a stable one) the usual analysis of a single turning point in terms of Bessel functions cannot be simply repeated (Langer, 1959) at each of the two turning points. A solution should rather be sought in terms of Weber functions allowing to cross the two turning points at once.A considerable amount of literature exists on the subject and a very recent paper by Olver (1975) opens the way to straightforward applications. They may bring to our attention, at least in special cases, unstable solutions which,in the above case, may have large and comparable amplitudes in the two unstable regions with a relatively minor reduction of this amplitude

https://doi.org/10.1017/S0252921100112333 Published online by Cambridge University Press

across the stable region especially if the latter is fairly narrow with a density gradient only slightly subadiabatic. Such modes could be particularly efficient in mixing the whole star. This conjecture is somewhat supported by some simplified or numerical investigations. For instance, Goosens and Smeyers (1974) have found more or less "accidental resonances" between some of the stable g modes of two stable regions separated by an unstable one (just the opposite of the case considered above) giving rise to a stable mode with large amplitudes in both stable regions decreasing only moderately in between in the unstable region.

Other examples have been treated by students in Liège which lead to similar conclusions. Consider for instance, the case of a heterogeneous incompressible model composed of superposed layers of different densities presenting two unstable discontinuities, $(\rho_{in} - \rho_{ext})<0$, separated by a stable one. The behaviour of the eigenfunctions associated with the two negative eigenvalues can depend drastically on the closeness of these eigenvalues. In general, i.e. as long as those two σ^2 are not very close, each of the eigenfunctions has a single maximum at the unstable interface with which it is associated. However when the parameters of the discontinuities are varied, one of these solutions may acquire a secondary strong maximum at the other discontinuity, the minimum between the two remaining appreciable when this eigenvalue becomes very close to that corresponding to the other discontinuity.

3. THE EFFECTS OF THE DYNAMICAL MOTIONS AND THE REDUCTION OF THE SUPERADIABATIC GRADIENT

Many of the examples where dynamically unstable modes have been found are artificial because the superadiabaticity (A) has been fixed a priori at a much larger value than is ever likely to occur in stars. In such cases, the time-scales of the growing motions which are proportional to $(A)^{-1/2}$ are rather short. Of course in the end, these violent motions lead to the establishment of a convective zone through which A is reduced to a very small value just sufficient to allow the residual energy no longer transferred by radiation to be carried by convective currents. This implies important readjustments in the internal structure of the star including transfer of mass to deeper layers and various feed-back effects which may increase considerably the extent of the convection zone with respect to the initial superadiabatic region. For instance, in going from an homogeneous compressible sphere to a polytrope $n = 3/2 (\Gamma_1 = 5/3)$, the energy released, if the mass and the radius are those of the sun, is of the order of $0.1\ GM_\odot^2/R_\odot = 5.10^{47}$ ergs in a short time of the order of one hour. Of course this is an extreme case, but even if the energy release is reduced by a factor 10^{10} and the duration increased by a few orders of magnitude, it would still remain a fairly spectacular phenomenon which was considered at one time (Biermann, 1939; Schatzman, 1946) significant for the interpretation of novae.

In any case, I suppose that there are no serious doubts that the readjustments contemplated above would lead in the end to the same model as the one that could be built a priori using the usual method of having a convective adiabatic zone initiated at the point where the radiative gradient becomes exactly equal to the adiabatic one ($A = 0$).

In fact things are a little more subtle as the reduction of the superadiabaticity must proceed only so far that the subsisting excess provides the necessary buoyancy force to balance the energy dissipated by viscosity and conduction (radiative or otherwise). Of course, when A and $|\sigma^2|$ are still large, the effects of these dissi-

https://doi.org/10.1017/S0252921100112333 Published online by Cambridge University Press

pation terms could be evaluated by a perturbation method (similar to that used for vibrational stability when $\sigma^2>0$) yielding a damping coefficient σ' correcting the adiabatic time dependence by a factor $e^{-\sigma' t}$.

But one knows how cumbersome (cf. Ledoux 1974), the expression of σ' is. To illustrate the situation, it seems better to revert to a simple case as, for instance, the plane layer with constant coefficients. Let κ and η represent respectively the thermometric conductivity and the kinematic coefficient of viscosity. Assuming a time dependence e^{st}, one finds for the g modes, if A is positive and larger than κ, η or $\kappa+\eta$

$$s = \pm \left(\frac{k_H^2}{k^2} gA\right)^{1/2} - \left(\frac{\kappa+\eta}{2}\right) k^2 \tag{18}$$

where k_H and k_z ($k^2 = k_H^2 + k_z^2$) are respectively the horizontal and the vertical wave number. In this case, dissipation simply hinders a little convection but does not affect it very much.

On the other hand, when convection is established, the effects of the dissipation forces are of the same order as those of the residual buoyancy and the above formula is no longer significant. In that case and if one includes the rate of energy generation ε (significant in the deep interior), one gets an equation equivalent to that of Defouw (1970) with

$$L_T = -\frac{\nu\varepsilon}{T}, \quad L_\rho = -\frac{\varepsilon}{\rho}$$

where ν represents the sensitivity of ε to T, $\nu = (d\log\varepsilon)_\rho / d\log T$. After separating a secular root $s_3 = -\eta k^2$, the dispersion relation for g modes, gives solutions

https://doi.org/10.1017/S0252921100112333 Published online by Cambridge University Press

$$s = -\frac{1}{2}\left\{ (\eta + \kappa)k^2 - \frac{\varepsilon}{C_p T}(\nu - 1) \right\}$$

$$\pm\left\{\frac{1}{4}\left[(\kappa - \eta)k^2 - \frac{\varepsilon}{C_p T}(\nu - 1)\right]^2 + \frac{k_H^2}{k^2} gA\right\}^{1/2} \tag{19}$$

which for vanishing ε reduces to the ordinary Rayleigh solution of the Bénard problem

$$s = -\frac{1}{2}(\eta + \kappa)k^2 \pm\left\{\frac{1}{4}(\kappa - \eta)^2 k^4 + \frac{k_H^2}{k^2} gA\right\}^{1/2} \tag{20}$$

which yields back (18) if $(k_H^2/k^2)gA >> \kappa^2 k^4$ and $\eta^2 k^4$

On the other hand, if we approach the marginal case ($s=0$), then $(k_H^2/k^2)gA$ must take the appropriate value to make the square root exactly equal to the first term, i.e.

$$\frac{k_H^2}{k^2} gA = \kappa\eta k^4 - \eta k^2 \frac{\varepsilon}{C_p T}(\nu - 1)$$

If ε is negligible, this condition becomes

$$\frac{gAd^4}{\kappa\eta} = \frac{k^6 d^4}{k_H^2}$$

where the depth d of the layer has been introduced. If $k_z = \pi/d$ and $a = k_H d$, one gets finally

$$\frac{gAd^4}{\kappa\eta} = \frac{(\pi^2 + a^2)^3}{a^2} = R$$

which is the usual value of the Rayleigh number. The energy generation could reduce somewhat the value of R. In this marginal case,

https://doi.org/10.1017/S0252921100112333 Published online by Cambridge University Press

the time-scale $\approx 1/s$ becomes infinite but a circulation sets in with a finite turn-over time.

Thus convection, although originating essentially in a dynamical instability (conservative terms) with a short time scale, initiates fast motions to modify the medium itself in such a way as to reduce gA finally to the same order as the dissipative (non conservative) forces.

As Defouw (1970) pointed out, the energy generation term has a destabilizing influence (cf.19) and if

$$\frac{\varepsilon}{C_p T}(\nu-1) > (\eta+\kappa)k^2$$

it contributes directly to the instability. If $A>0$, this effect simply reinforces the buoyancy.

On the other hand if $A<0$, and such that

$$-\frac{k_H^2}{k^2}\, gA > \frac{1}{4}\left\{(\kappa-\eta)k^2 - \frac{\varepsilon}{C_p T}(\nu-1)\right\}^2$$

the effect would correspond to a case of vibrational instability (or overstability) of stable g modes. But if

$$-\frac{k_H^2}{k^2}\, gA < \frac{1}{4}\left\{(\kappa-\eta)k^2 - \frac{\varepsilon}{C_p T}(\nu-1)\right\}^2$$

it seems that a growing non-oscillatory motion would arise which might lead to some kind of convection, although $A<0$.

REFERENCES

BIERMANN, L. 1939, Z. Astrophys., 18, 344

CHANDRASEKHAR, S. 1964, Astrophys. J., 139, 664

COWLING, T.G. 1941, Mon. Not. Roy. Astr. Soc., 101, 367

DEFOUW, R.J. 1970, Astrophys. J., 160, 659

DEUPREE, R.G. 1974, Astrophys.J., 194, 393

DEUPREE, R.G. 1975a, Astrophys.J., 198, 419

DEUPREE, R.G. 1975b, Astrophys.J., 201, 183

DEUPREE, R.G. 1976, Astrophys.J., 205, 286

GOOSENS, M. and SMEYERS, P. 1974, Astrophys. Space Sci., 26, 137

GRISVARD, P., SOUFFRIN, P. and ZERNER, M. 1972, Astron. and Astrophys., 17, 309

ISLAM, J.N. 1970, Mon. Not. Roy. Astr. Soc., 150, 237

KOVETZ, A. 1967, Z. Astrophys., 66, 446

LANGER, R.E. 1959, Trans. Amer. Math. Soc., 90, 113

LEBOVITZ, N.R. 1965a, Astrophys. J., 142, 229

LEBOVITZ, N.R. 1965b, Astrophys. J., 142, 1257

LEBOVITZ, N.R. 1966, Astrophys. J., 146, 946

LEDOUX, P. 1949, Contribution à l'étude de la structure interne des étoiles et de leur stabilité, Mém. Soc. Roy. Sci. Liège Coll. in -8°, 4°sér. T IX, Ch. III, sect. 4, 5, 6

LEDOUX, P. et SMEYERS, P. 1966, Compt. Rend. Acad. Sci. Paris, Sér.B., 262, 841

LEDOUX, P. 1974, in P.Ledoux et al (ed.) Stellar Instability and Evolution, IAU Symposium n°59, Part VI, p. 135

MOULTON, F.R. 1909, Astrophys. J., 29, 257

NORDLUND, Å. 1976, Astron. and Astrophys., 50, 23

OLVER, F.W.J. 1975, Phil. Trans. Roy. Soc. London A, 278, 137

OSAKI, Y. 1975, Publ. Astron. Soc. Japan, 27, 237

PEKERIS, C.L. 1938, Astrophys. J., 88, 189

SCHATZMAN, E. 1946, Ann. Astrophys., 9, 199

SCUFLAIRE, R. 1974, Astron. and Astrophys., 34, 449

SHAPLEY, H. 1914, Astrophys. J., 40, 448

SPIEGEL, E.A. 1971, Ann. Rev. Astron. Astrophys. 9, 323

SPIEGEL, E.A. 1972, Ann. Rev. Astron. Astrophys., 10, 261

TASSOUL M. and TASSOUL J.L. 1968, Ann. Astrophys., 31, 251

THOMSON, W. 1863, Phil. Trans. Roy. Soc. London, 153, 612

THORNE, K.S. 1966, Astrophys. J., 144, 201

https://doi.org/10.1017/S0252921100112333 Published online by Cambridge University Press