HISTORICAL REMINISCENCES OF THE ORIGINS OF STELLAR CONVECTION THEORY

(1930 - 1945)

Ludwig Biermann
Max-Planck-Institut für Physik und Astrophysik
Munich, Germany

To set the stage for the report to follow+), let me start with a quotation from A.S. Eddington's classical work "The Internal Constitution of the Stars", published exactly half a century ago. Eddington wrote: "We shall not enter further into the historic problem of convective equilibrium since modern researches show that the hypothesis is untenable. In stellar conditions the main process of transfer of heat is by radiation and other modes of transfer may be neglected." (paragraph 69, p. 98).

Only 19 years earlier Robert Emden had outlined for the sun a very different picture; in paragraph 4, chapter 18, of "Gaskugeln" - also a classical treatise which influenced research well into the thirties - he had stated with equal conviction that the energy radiated into space from the photosphere could be brought there almost exclusively only by convection, and that the granulation depicts the cross section of the ascending hotter and the descending darker currents. Emden's manuscript had been read in the proof stage by his brother-in-law Karl Schwarzschild, who advised Emden on a number of points. It seems worth noting that Schwarzschild's somewhat earlier work on the radiative equilibrium of the sun's atmosphere was discussed in Emden's monograph (in chapter 16, par. 13)++).

+) The text includes a number of points, which actually came up in the later discussions during the colloquium or, in one case, during the IAU assembly in Grenoble. The author is indebted to many colleagues, in particular to D. Gough, L. Mestel, M. Schwarzschild and N. Weiss for important comments.

++) Whether Schwarzschild, who had proven that radiative transport prevailed in the sun's photospheric layer and had formulated the quantitative criterion for the stability of such layers, was in complete agreement with Emden, is not quite clear, though Emden's text does convey this impression.

https://doi.org/10.1017/S0252921100112291 Published online by Cambridge University Press

Part of the change of the scientific argument during the intervening time (1907-1926) was of course due to Jeans's and Lindemann and Newall's discovery that at the pressures and temperatures prevailing in the stellar interiors, all the molecules and atoms would be broken up into almost bare atomic nuclei and electrons and that, as a consequence, radiative transport was recognized to be most efficient, in Eddington's scheme of 1926 almost too efficient. In retrospect it looks however as if Eddington, in laying the foundations of the theory of radiative equilibrium in stars, did apparently not fully appreciate the power of even very slow turbulent convection in transporting energy in stellar interiors; that he neglected the possible influence of the surface conditions was perhaps related to the fact that they were, at the time when Eddington wrote his book, essentially inapplicable to most stars.+)

This review will be concerned mainly with three developments, which took place between 1930 and 1946: 1) the application to stellar interiors of the convective transport equation developed in hydrodynamics, which led to the proof that the adiabatic temperature gradient is, in the case of thermal instability, a very good approximation there; 2) the influence of the surface boundary condition in determining the extent of outer convection zones, which in certain circumstances may comprise the whole star; 3) the introduction of the scale height as a measure of the mixing length used in the transport equation. Their connection with some further developments which may come up during the present conference can be sketched only very briefly.

To begin with item 1, we note first that observations and measurements of transport processes by non-stationary ("turbulent") mass motions in the earth' atmosphere and in the oceans had led, between 1915 and 1925, to a reasonably successful theoretical scheme. G.I. Taylor (1915) and W. Schmidt (1917) noted++) that the quotient of the

+) Concerning both points it is instructive to reread his discussion of the point source model (ICS, § 91). It should be added, however, that Eddington (as far as the present author was able to find out) fully accepted the change of position which occurred during the period under review in this report.
++) Apparently, as a result of the war conditions, independently, as can be judged from their papers of 1915 and 1917.

flux of some quantity - like the momentum, the salinity or heat - and the gradient of the content per cm^3 or per gram of the same quantity led to consistent values of this coefficient, for which W. Schmidt introduced the term "Austausch" (gr $cm^{-1}sec^{-1}$). In the special case of heat flow ("Scheinleitung") the excess of the actual temperature gradient over its adiabatic value, $\Delta\nabla T$, multiplied with the specific heat for constant pressure, has to be used, because only small pressure differences arise if the Mach number is <<1, as is usually true in the earth's atmosphere. For interpreting the observed values of the "Austausch", concepts developed by G.I. Taylor and by L. Prandtl (1925), in particular the "mixing length", were most useful. (These had been introduced first for the case of dynamical instability.) These concepts relied to some extent of considerations analogous to those of kinetic theory. The analogue of the molecule of kinetic theory is an element of the fluid, which (having detached from its surroundings as a consequence of the given instability due either to a super-adiabatic temperature gradient or to the dynamical situation, for instance shear) moves as a whole over some distance until it mixes again with the surrounding fluid. In Prandtl's original thought the mixing length ℓ was determined by the geometry of the situation, e.g. the distance from the nearest boundary or the diameter of the unstable region, and this was used in the first application to stellar interiors. The somewhat later idea, to link the mixing length with the scale height, will be discussed below.

To determine the velocity it was considered that pressure equilibrium is only slightly disturbed, whereas - the flow of heat by conduction or radiation being usually much slower - the temperature of an element of the fluid is determined by the adiabatic gradient, such that in case of thermal instability a rising element is less dense and hotter than the surrounding fluid, and as a consequence is accelerated. These considerations led to an expression for the velocity (v) of the moving elements, which can be written in the form (g = acceleration of gravity)

$$v^2 \approx g\ell \cdot \frac{\Delta T}{T} \approx g\ell^2 \frac{\Delta\nabla T}{T} \quad .$$

For the convective transport of heat (H_K = erg/cm^{2}sec, c_p = specific heat per gr for constant pressure, ρ = density) we write

$$H_K \approx c_p \cdot \rho \ell v \cdot \Delta\nabla T$$

$$\approx pv \cdot \frac{v^2}{g\ell} \quad (\frac{\mu c_p}{R^*})$$

(p - pressure $=(R^*/\mu)\rho T$; R^* - gas constant, μ - mean molecular weight), suppressing again a constant of order unity.

In meteorology the largest scale for which the transport equation has been used with at least qualitative success (by A. Defant, see W. Schmidt 1925), is the meridional transport of heat from the equatorial to the polar regions, which raises the average temperatures of our own latitudes quite noticeably, the meridional heat flux exceeding the "solar constant" by roughly two powers of ten (as it must). In this case the elements are low pressure systems which, due to their rotation, have a certain stability; it happens, that the mixing length and the pressure, but of course not the temperature and the density, are comparable to those in the outermost layers of the sun's hydrogen convection zone.

The application to stellar interiors was contained in a Göttingen thesis of 1932 (L. Biermann 1933), which owed a great deal to Prandtl's advice. For the convective zone around the centre, due to the conversion of hydrogen into helium (supposed to be highly temperature sensitive as is true for the C-N cycle), the mixing length was taken to be of the order of 10.000 km, and the flux to be transported assumed to be of the order of 10^{12}erg/cm^{2}sec, about 15 times that on the sun's surface. For the third power of the velocity a combination of the equations given above leads to

$$v^3 \approx \frac{H_K g\ell}{p} \quad (\frac{R^*}{\mu c_p})$$

which was found to be or order 10^{11}(cm/sec)^3. This then leads to

$$|\Delta\nabla T| \approx 10^{-8} \quad \text{or} \quad \approx 10^{-5}|\nabla T| \;.$$

Though there is some incertainty about the coefficients of order unity contained in the equations given above, it is clear that the

https://doi.org/10.1017/S0252921100112291 Published online by Cambridge University Press

result last stated, that a relatively very small excess over the adiabatic gradient is sufficient to carry all the flux, is quite insensitive to errors made by the phenomenological theory used here. Two years later Thomas G. Cowling, who used a slightly different approach and formalism, recovered the same result. Since 1933/34 it has become standard practice to use, in theoretical models of star's interiors, the adiabatic temperature gradient whenever that required for radiative transport exceeds the adiabatic one+). In this form the stability criterion was formulated first by Karl Schwarzschild for the solar atmosphere (1906); in meteorology an equivalent criterion had been in use already many years earlier.

Concerning item 2, we note first, that during the years 1934-38, an attempt was made to explore the question whether partially or wholly convective stellar models (not only such with a central convection zone) lead to a more complex picture of the overall constitution of the stars than the one given by Eddington (as was finally found to be the case). Such models would result from a higher luminosity than the radiative ones, for the same radius and opacity if the luminosity could be regarded as an (effectively) free parameter. Near the surface two circumstances, the second of which was not at once fully appreciated have to be taken into account: the existence, in all stars with not too high surface temperature, of convective zones due to the partial ionization of hydrogen and helium, which had first been investigated by Unsöld in 1930, and (second) the necessity of using in the photosphere of every star the radiative transport equation together with that of hydrostatic equilibrium, which means that the pressure must be of order g/κ, κ being the opacity ($gr^{-1}cm^2$) of the photosphereic layers - a relation which of course could be written down with better accuracy. The importance of this boundary condition for the case under discussion (L. Biermann 1935) was emphasized by T.G. Cowling in 1936, but its application was held back for some years by the (then) poor knowledge of the value of the photospheric opacity for all stars later than spectral class A.

+) The first such model being Cowling's point source model of 1935.

https://doi.org/10.1017/S0252921100112291 Published online by Cambridge University Press

Shortly before the discovery by Rupert Wildt that the negative hydrogen ion is the main source of the photospheric opacity in such stars, an attempt was made (L. Biermann 1938) to use approximate values for the photospheric opacity based on work of Pannekoek which fortunately led to approximately correct answers. It was found (a result which was soon confirmed on the basis of Wildt's pioneering work of 1939, and subsequent work of others) that in the surface of the sun and similar stars (with photospheric pressures of the order of 10^5dyn/cm^2, such that the radiation pressure is only $\approx 10^{-5}$p) convection was likely to be an efficient mechanism of transport of heat not only in the main parts of the hydrogen convection zone, but also in its outer layers; as a consequence the adiabatic gradient should at least approximately be established in the hydrogen convection zone up to its outer boundary in the middle or deep photosphere.

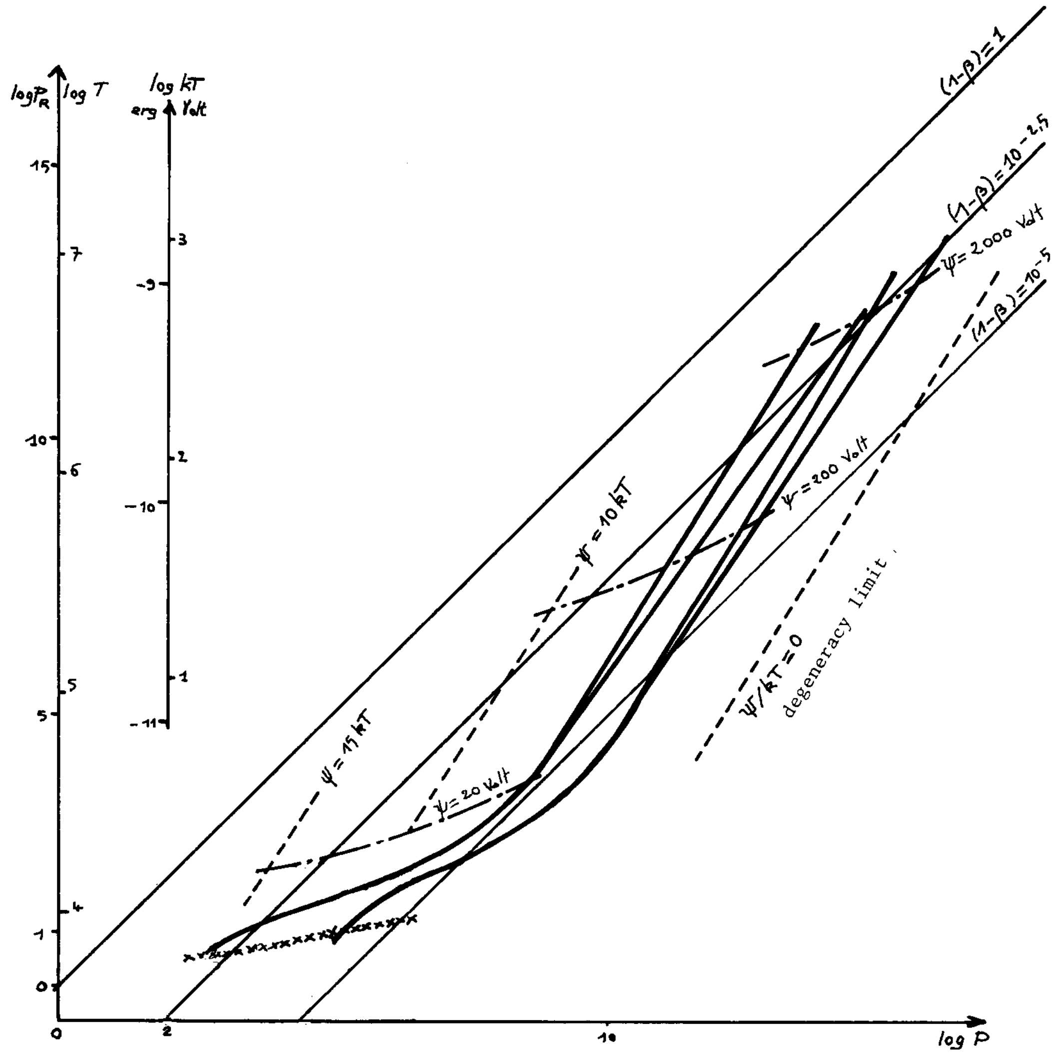
In order to illustrate the importance of this result let us look at a diagram taken from the author's paper of 1938. This diagram, with the logarithm of the total pressure ($P = p+p_R$) and of radiation pressure (p_R) as coordinates, shows besides the adiabats (full lines+)) the lines of constant ratio p_R to P (Eddington $1-\beta$, weak lines), furthermore the "dominant" ionization potential ψ++) and lines for fixed ratio of ψ to kT, of which the one for $\psi = 0$ corresponds to the limit of degeneracy ("Entartungsgrenze"). Near the bottom are shown the photospheric values of the pressure (for given effective temperature) for three values of the surface gravity, including that for the sun. The crossed line marked

$$\left(\frac{d \log p_R}{d \log P}\right)_{ad} = 1$$

indicates the zone of rapid increase (inwards) of the opacity and as a consequence of the radiative temperature gradient, and the incipient decrease of the adiabatic gradient (inwards) due to the additional degrees of freedom (using the terminology of the kinetic theory of the

+) The branching above $\log T \approx 5$ corresponds to then existing uncertainties regarding the chemical composition, especially the value of Z (in the terminology in use now).

++) For which the degree of ionization is $\approx 1/2$ according to Saha's formula.



https://doi.org/10.1017/S0252921100112291 Published online by Cambridge University Press

specific heat) resulting from the increasing degree of ionization. Above ψ = 50 eV, the ratio of the radiative gradient to the adiabatic gradient

$$\frac{(d \log p_R/d \log P)_R}{(d \log p_R/d \log P)_{ad}}$$

is essentially given by the value of 1-β, such that for the special solution corresponding to Eddington's model, for given mass and luminosity, and for larger values of 1-β, the radiative gradient is smaller than the adiabatic one and Schwarzschild's stability criterion is fulfilled. For smaller values of 1-β the radiative gradient increases, such that the instability limit is reached soon and the adiabatic gradient (d log p_R/d log P $\approx$ 8/5 for stars of about the sun's mass, for which 1-β <<1) is the smaller one and convection prevails.+)

It is thus seen+) that the surface boundary condition effectively results in such a relation between p_R and P that the total pressure increases inwards at first considerably faster than the radiation pressure; in the regions with temperatures around some 10^4 degrees, 1-β has a minimum and the radiative gradient is much larger than the adiabatic one, the opacity being approximately proportional $(1-\beta)^{-1}$ according to Kramers' law (in that temperature region actually still larger), whereas the adiabatic gradient is approximately constant ($\approx$ 8/5). With the increase of 1-β along the adiabates inwards, the stability limit is gradually approached and in the case of the sun finally reached at T $\gtrsim 10^6$ (with modern values of the chemical composition Eddington's model would show 1-$\beta \approx 10^{-3}$ for the sun). It is therefore clear that in the sun and in similar stars the outer convection zone must comprise a substantial number of scale heights until the radiative temperature gradient decreases below the adiabatic one, such that the inner boundary of the hydrogen convection zone should be at a temperature T $\gtrsim 10^6$ and a depth of $\gtrsim$ 100 000 km. This result has been confirmed by the much more accurate computations of recent years. Dwarf stars of later spectral class should have still deeper convection zones,

---

+) The presentation attempts to retrace the steps, by which the complete stellar models with deep outer convection zones and the fully convective models were actually found. For a more detailed review see L. Biermann 1945.

https://doi.org/10.1017/S0252921100112291 Published online by Cambridge University Press

and it was proposed in 1938 that late type giants might even be fully convective, a result which was recovered many years later by Hayashi.

In retrospect, it is easily seen why all these possibilities had not been noticed earlier: that convective transport could be efficient up to almost photospheric layers, was a rather remote possibility on the background of the earlier theory - though not on that of Emden - and a quantitative discussion of the power of convective transport was hindered by the lack of reliable knowledge of the photospheric pressure (which had been highly underestimated before 1938/39).

The largest uncertainty in transferring the mixing length theory to astrophysics - our item 3 - is evidently connected with the value of ℓ to be used, under the different circumstances. For the work done in 1938 described above the observed size of the solar granulation had been taken as a guide; this choice+) which fortunately did not introduce serious errors was until 1943 replaced by the answer which has been the basis of all subsequent work, and which is to use the local scale height, defined either by the density gradient or that of the pressure, as measure of ℓ, such that ℓ is given:

$$\ell \approx \frac{1}{|\nabla \ell n \rho|} \text{ or } \frac{1}{|\nabla \ell n P|} .$$

Since there should be a nondimensional factor of order unity, which requires separate discussion, the two expressions are under most circumstances equivalent. The idea behind this choice is that in any case the largest elements should travel farthest and reach the highest velocity but that an element of the fluid, after having travelled over a density scale height, should have changed its shape to such an extent that it is likely to break up into smaller fragments and to mix with the surrounding fluid. This approximation is of course precisely in the spirit of Prandtl's original ideas on the subject, and had most probably been discussed with him. On a slightly different background, an equivalent proposal had been made by E. Öpik already in 1938.++)

+) An at least approximate determination of the size distribution function of the solar granulation became possible recently (J.W. Harvey, M. Schwarzschild, 1975), whereas for red giants the observational situation is still less clear (M. Schwarzschild, 1975).

++) The work reported here under items 1) and 2) had remained unknown to Öpik until his work of 1938 was completed, cf. his "Note added in proof" (Öpik 1938). For a more recent review of the general problem see M. Schwarzschild, 1961).

To conclude these reminiscences I would first like to mention that the results discussed under item 2), led for the sun to a proposal concerning the old question, why a sunspot is dark, the answer being, that the strong magnetic fields observed in the umbra should inhibit the convective transport in the layers underneath the spot[+)]; it has been pursued by a number of authors up to the present and may come up again at this colloquium.

Of the various formalisms suggested to improve upon the one given above for the heat transport, the scheme proposed by E. Böhm-Vitense 1953, 1958 became the most widely used. Much more recently R. Ulrich 1970 has proposed still further refinements, which are particularly useful at relatively low densities. All attempts to determine the exact relationship between the mixing length and the scale height (and/or other parameters) have not been really satisfactory so far, though comparisons with observational data on the integral properties of stars (including their chemical composition) and on the position of the instability strip in the Herzsprung-Russell Diagram have been used with some success; only for the sun its known age provides an independent parameter and thereby a check that the mixing length theory leads to essentially reliable results (D.O. Gough, N.O. Weiss 1976). It seems that only a deductive theory of stellar convection would offer the chance to go beyond the present essentially phenomenological approach used hitherto; at least one contribution at the present colloquium will, I understand, deal with this problem.

+) L. Biermann 1941.

https://doi.org/10.1017/S0252921100112291 Published online by Cambridge University Press

References

Biermann, L., 1933, Z. Astrophys. 5, 117,

Biermann, L., 1935, Astr. Nachr. 257, 269.

Biermann, L., 1938, Astr. Nachr. 264, 395.

Biermann, L., 1941, Vierteljahresschrift Astron. Ges. 76, 194.

Biermann, L., 1943, Z. Astroph. 22, 244.

Böhm-Vitense, E., 1958, Z. Astrophys. 46, 108.

Cowling, T.G., 1935, Mon.Not.R.astr.Soc. 96, 18.

Cowling, T.G., 1936, Astron. Nachr. 258, 133.

Eddington, A.S., 1926, Internal Constitution of Stars

Emden, R., 1907, "Gaskugeln"

Gough, D., Weiss, N.O., 1976, Mon.Not.R.astr.Soc. 176, 589.

Harvey, J.W., Schwarzschild, M., 1975, Ap.J. 196, 221.

Öpik, E.J., 1938, Publ. Obs. astr. Univ. Tartu 30, No. 3.

Prandtl, L., 1925, Zeitschr. f. angew. Math. u. Mech. 5, 136.

Prandtl, L., 1932, Beitr. Physik freier Atmosphäre 19, 188.

Schmidt, W., 1917, "Der Massenaustausch bei der ungeordneten Strömung in freier Luft und seinen Folgen", Kaiserl. Akad. d. Wiss., Wien.

Schmidt, W., 1925, "Der Massenaustausch in freier Luft und verwandte Erscheinungen", Verlag von Henri Grand, Hamburg.

Schwarzschild, K., 1906, Nachr. Königl. Ges. d. Wiss., Göttingen, No.1.

Schwarzschild, M., 1961, Ap. J. 134, 1.

Schwarzschild, M., 1975, Ap. J. 195, 137.

Taylor, G.I., 1915, Phil. Trans. R. Soc. Lond. A, 215, 1.

Taylor, G.I., 1932, Proc. R. Soc. Lond. A, 135, 685.

Ulrich, R., 1970, Astroph. & Space Science 7, 71.

Ulrich, R., 1970, Astroph. & Space Science 7, 183.

Unsöld, A., 1930, Z. f. Astroph. 1, 138.

Vitense, E., 1953, Z. Astrophys. 32, 135.

Wild, R., 1939, Astroph. J. 90, 611.