Section 1

The Equations of Stellar structure

1.1 Introduction: motivation

In this part of the course, we consider aspects of the internal operation of (principally) single stars: their structure and evolution. Our overarching aim in this is to interpret observations such as the Hertzsprung–Russell diagrams shown in Fig. 1.1

For the present purposes, we use a working definition of a star as an isolated body that is bound by self-gravity, and which radiates energy supplied by an internal source. Self-gravity ensures that the star is approximately spherical (rotation introduces centrifugal forces which, for sufficiently fast rotation, may introduce distortions); the internal source of energy is nuclear fusion for most of the stellar lifetime (although for, e.g., white dwarfs, stored thermal energy is responsible for the observed luminosity).

The essence of stellar structure is the competition between the force of gravity, which always wants to make a star collapse, and the outward force of pressure. For almost the entire lifetime of a star, these forces are in balance; the star is in (or very close to) hydrostatic equilibrium, but as internal energy is released, the internal composition, and hence structure, must evolve. Thus 'stellar structure' and 'evolution' are intimately linked.

1.2 Review: Basic Equations of Stellar Structure

For reference, we remind ourselves of the basic equations of stellar structure, introduced in PHAS 2112 and Dr. Zane's notes:

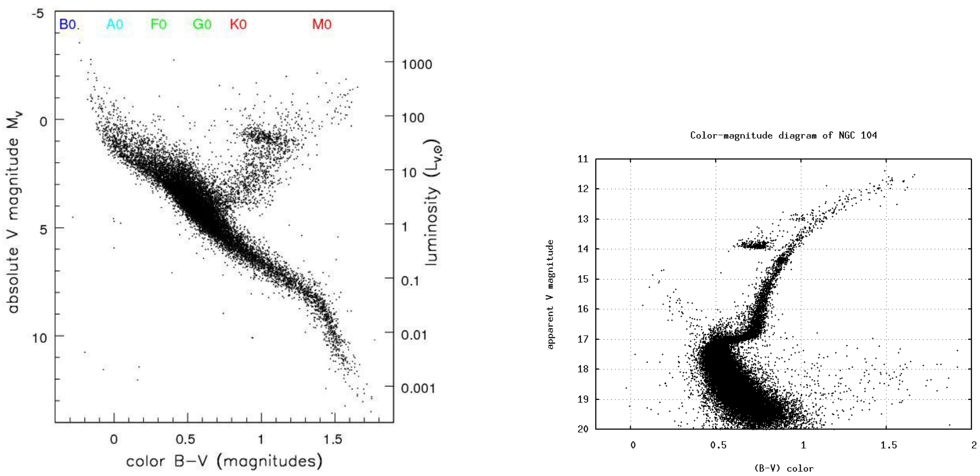

Figure 1.1: Hertzsprung–Russell (colour–magnitude) diagrams. Left, *Hipparcos* volume-limited sample (stars of different ages); right, *HST* observations of the globular cluster 47 Tuc (coeval sample).

1.2.1 Hydrostatic equilibrium; equations of state

$$\frac{\mathrm{d}P(r)}{\mathrm{d}r} = \frac{-Gm(r)\rho(r)}{r^2} = -\rho(r)\, g(r) \tag{1.1}$$

or

$$\frac{\mathrm{d}P(r)}{\mathrm{d}r} + \rho(r)\, g(r) = 0$$

[In the supplementary 2112 notes, eqtn. 12.1; this numbering may well differ from that in use when you took in PHAS 2112].

The principal sources of pressure throughout a normal star are gas pressure, and radiation pressure.[1] We will take the corresponding equations of state to be, in general,

$$\begin{aligned} P_{\mathrm{G}} &= nkT; \\ &= (\rho kT)/(\mu m(\mathrm{H})) \qquad (1.2) \\ P_{\mathrm{R}} &= \frac{1}{3} a T^4 \qquad (1.3) \end{aligned}$$

for number density n at temperature T, density ρ; μ is the mean molecular weight, and $m(\mathrm{H})$ the hydrogen mass; a is the radiation constant, $a = 4\sigma/c$; with σ the Stefan-Boltzmann constant, and k Boltzmann's constant.

[1] Electron degeneracy pressure is important in white dwarfs, and neutron degeneracy pressure in neutron stars.

1.2.2 Mass Continuity

$$\frac{\mathrm{d}m}{\mathrm{d}r} = 4\pi r^2 \rho(r); \tag{1.4}$$

[PHAS 2112 eqtn. 12.3] this form holds in the case of hydrostatic equilibrium.

1.2.3 Energy continuity

$$\frac{\mathrm{d}L}{\mathrm{d}r} = 4\pi r^2 \rho(r)\epsilon(r) \tag{1.5}$$

[PHAS 2112 eqtn. 12.7].

where

- r is radial distance measured from the centre of the star
- $P(r)$ is the total pressure at radius r
- $\rho(r)$ is the density at radius r
- $g(r)$ is the gravitational acceleration at radius r
- $m(r)$ is the mass contained with radius r
- $L(r)$ is the total energy transported through a spherical surface at radius r
- $\epsilon(r)$ is the energy generation rate per unit mass at radius r

The stellar radius is R_*, the stellar mass is $M_* \equiv m(R_*)$, and the emergent luminosity $L_* \equiv L(R_*)$ (dominated by radiation at the stellar surface).

1.2.4 Energy transport

We suppose that the Sun is hotter inside than outside, so there must be an energy flow. We are familiar with three basic mechanisms of energy transport:
– radiation
– convection
– conduction

In the context of stellar astrophysics, conduction is important only under the degenerate conditions found in white dwarfs and neutron stars (since gases in general are poor conductors). For ‘normal’ stars, the key processes transporting energy are radiation and convection.

Radiative transport: Energy is transported by photons. In stellar interiors the opacities are high, and the mean free path correspondingly low – about 1 mm in the case of the Sun (Section 14.4). In this sense, the radiation doesn't *flow* outwards, but rather *diffuses* outwards.

Convective transport: If the radiation is unable to escape a layer at a rate that matches the energy input, then 'something's got to give'. What gives is the static nature of the layer: convection is initiated and starts to transport energy. This suggests that hydro*static* equilibrium breaks down, but the dynamical timescale is short compared to the flow timescale (Section 14), so in practice HSE continues to be an excellent approximation.

The nett energy flux is, under most circumstances, simply the sum of radiative and convective terms,

$$L(r) = L_{\rm rad}(r) + L_{\rm cnv}(r)$$

1.3 Radiative energy transport in stellar interiors

1.3.1 The equation of radiative transfer in stellar interiors

In optically thick environments – in particular, stellar interiors – radiation is often the most important transport mechanism, but for large opacities the radiant energy doesn't flow directly outwards; instead, it *diffuses* slowly outwards.

To express this transport quantitatively, the same general principles may be applied as led to the equation of radiative transfer in plane-parallel stellar atmospheres

$$\mu \frac{{\rm d}I_\nu}{{\rm d}\tau_\nu} = I_\nu - S_\nu. \tag{1.6}$$

[PHAS 2112 eqtn. (5.6), and Dr. Zane's lectures]; there is no azimuthal dependence of the radiation field, and the photon mean free path is (very) short compared to the radius. The conditions appropriate to 'local thermodynamic equilibrium' (LTE; PHAS 2112 Sec. 13.1) apply, and so the radiation field is very well approximated by black-body radiation.

Box 1.1. It may not be immediately obvious that the radiation field in stellar interiors is, essentially, isotropic; after all, outside the energy-generating core, the full stellar luminosity is transmitted across any spherical surface of radius r. However, if this flux is small compared to the local mean intensity, then isotropy is justified.

The flux at an interior radius r (outside the energy-generating core) must equal the flux at R (the surface); that is,

$$\pi F = \sigma T_{\text{eff}}^4 \frac{R^2}{r^2}$$

while the mean intensity is

$$J_\nu(r) \simeq B_\nu(T(r)) = \sigma T^4(r).$$

Their ratio is

$$\frac{F}{J} = \left(\frac{T_{\text{eff}}}{T(r)}\right)^4 \left(\frac{R}{r}\right)^2.$$

Temperature rises rapidly below the surface of stars, so this ratio is always small; for example, in the Sun, $T(r) \simeq 3.85$ MK at $r = 0.9\mathrm{R}_\odot$, whence $F/J \simeq 10^{-11}$. That is, the radiation field is isotropic to better than 1 part in 10^{11}.

Equivalently, the temperature gradient from the centre of the Sun (for example) to the surface is

$$\frac{\Delta T}{\Delta r} = \frac{T_{\text{c}} - T_{\text{eff}}}{R_\odot} \simeq 10^{-2} \text{ K m}^{-1} \tag{1.7}$$

The photon mean free path is $\ell = 1/\kappa_\nu \simeq 1$ mm (from detailed models), so the temperature change over this distance is of order 10^{-5} K. The radiant energy density is $U = aT^4$, so the relative anisotropy $\Delta U/U = 4\Delta T/T \simeq 10^{-11}$ at 10^6 K.

Although the anisotropy is very small, the nett outflow is large – in fact, equal to the stellar luminosity.

We recall that, in general, the intensity I_ν is direction-dependent; i.e., is $I_\nu(\theta, \phi)$ (although the explicit angular dependence is generally dropped for economy of nomenclature). Multiplying eqtn. (1.6) by $\mu \equiv \cos\theta$ and integrating over solid angle, using $\mathrm{d}\Omega = \sin\theta \,\mathrm{d}\theta \,\mathrm{d}\phi = \mathrm{d}\mu \,\mathrm{d}\phi$, then

$$\frac{\mathrm{d}}{\mathrm{d}\tau_\nu} \int_0^{2\pi} \int_{-1}^{+1} \mu^2 I_\nu(\mu, \phi) \,\mathrm{d}\mu \,\mathrm{d}\phi = \int_0^{2\pi} \int_{-1}^{+1} \mu I_\nu(\mu, \phi) \,\mathrm{d}\mu \,\mathrm{d}\phi - \int_0^{2\pi} \int_{-1}^{+1} \mu S_\nu(\mu, \phi) \,\mathrm{d}\mu \,\mathrm{d}\phi;$$

The radiation field is axially symmetric, so

$$\frac{\mathrm{d}}{\mathrm{d}\tau_\nu} \int_{-1}^{+1} \mu^2 I_\nu(\mu) \,\mathrm{d}\mu = \int_{-1}^{+1} \mu I_\nu(\mu) \,\mathrm{d}\mu - \int_{-1}^{+1} \mu S_\nu(\mu) \,\mathrm{d}\mu. \tag{1.8}$$

The first two terms should be familiar as moments of the radiation field,

$$
\begin{aligned}
F_\nu &= 4\pi H_\nu \\
&= 2\pi \int_{-1}^{+1} I_\nu(\mu)\mu\,\mathrm{d}\mu.
\end{aligned}
\tag{1.9}
$$

[where H_ν is the Eddington flux, or first-order moment of the radiation field; PHAS 2112 eqtn. (3.9)], and

$$
K_\nu = \frac{1}{2}\int_{-1}^{+1} I_\nu(\mu)\mu^2\,\mathrm{d}\mu
\tag{1.10}
$$

[the second-order moment, or K integral; PHAS 2112 eqtn. (3.14)]. Since the radiation field is locally isotropic to a very good approximation we can take I_ν out of this integral, so

$$
\begin{aligned}
K_\nu &= \frac{1}{2}\frac{\mu^3}{3}I_\nu\Big|_{-1}^{+1} \\
&= \frac{1}{3}I_\nu \qquad \left[\equiv \frac{1}{3}J_\nu \text{ for isotropy}\right]
\end{aligned}
\tag{1.11}
$$

[PHAS 2112 eqtn. (3.15)].

Using these equations for the first two terms in eqtn. (1.8), and supposing that the emissivity has no preferred direction (as is true to an excellent aproximation in stellar interiors; Box 1.1) so that the source function is isotropic (and so the final term is zero), we obtain

$$
\frac{\mathrm{d}K_\nu}{\mathrm{d}\tau_\nu} = \frac{F_\nu}{4\pi}
$$

or, from eqtn. (1.11),

$$
\frac{1}{3}\frac{\mathrm{d}I_\nu}{\mathrm{d}\tau_\nu} = \frac{F_\nu}{4\pi}.
$$

In LTE we may set $I_\nu = B_\nu(T)$, the Planck function; and $\mathrm{d}\tau_\nu = -k_\nu\,\mathrm{d}r$ (where again the minus arises because the optical depth is measured inwards, and decreases with increasing r). Making these substitutions, and integrating over frequency,

$$
\int_0^\infty F_\nu\,\mathrm{d}\nu = -\frac{4\pi}{3}\int_0^\infty \frac{1}{k_\nu}\frac{\mathrm{d}B_\nu(T)}{\mathrm{d}T}\frac{\mathrm{d}T}{\mathrm{d}r}\mathrm{d}\nu
\tag{1.12}
$$

To simplify this further, we introduce the *Rosseland mean opacity,* $\overline{k}_\mathrm{R}$ $(= \overline{\kappa}_\mathrm{R}\rho)$,[2,3] defined by

$$
\frac{1}{\overline{k}_\mathrm{R}}\int_0^\infty \frac{\mathrm{d}B_\nu(T)}{\mathrm{d}T}\mathrm{d}\nu = \int_0^\infty \frac{1}{k_\nu}\frac{\mathrm{d}B_\nu(T)}{\mathrm{d}T}\mathrm{d}\nu.
$$

[2] Recall that opacity may be expressed in several ways, most commonly as 'per unit mass' or 'per unit volume'. We use k to denote opacity per unit volume, and κ where reference is made to opacity per unit mass; clearly, then, $k = \kappa\rho$.

[3] The Rosseland mean opacity represents the harmonic mean of k_ν, weighted by $\mathrm{d}B_\nu(T)/\mathrm{d}T$. This weighting factor is small for very low and very high frequencies, and peaks at $\nu_\mathrm{p} = 4kT/h$.

Recalling that

$$\int_0^\infty \pi B_\nu \, \mathrm{d}\nu = \sigma T^4 \qquad \text{(PHAS 2112, eqtn. (3.21))}$$

we also have

$$\begin{aligned}\int_0^\infty \frac{\mathrm{d}B_\nu(T)}{\mathrm{d}T}\mathrm{d}\nu &= \frac{\mathrm{d}}{\mathrm{d}T}\int_0^\infty B_\nu(T)\mathrm{d}\nu \\ &= \frac{4\sigma T^3}{\pi}\end{aligned}$$

(at given T) so that eqtn. (1.12) can be written as

$$\int_0^\infty F_\nu \, \mathrm{d}\nu = -\frac{4\pi}{3}\frac{1}{\bar{k}_\mathrm{R}}\frac{\mathrm{d}T}{\mathrm{d}r}\frac{acT^3}{\pi} \tag{1.13}$$

where a is the radiation constant, $4\sigma/c$; that is, the total radiant energy flux is

$$F = -\frac{4\pi}{3}\frac{1}{\bar{k}_\mathrm{R}}\frac{\mathrm{d}T}{\mathrm{d}r}\frac{acT^3}{\pi} \tag{1.14}$$

[This shows that radiative diffusion is completely analogous to conduction;

$$F \propto \frac{\mathrm{d}T}{\mathrm{d}r},$$

which is equivalent to Fourier's law of thermal conduction.]

This is our adopted form for the transport of radiative flux. It may be applied in environments where the photon mean free path is short compared to the scales over which physical parameters (notably temperature) change; it therefore becomes inappropriate as the stellar surface is approached, where a more detailed approach to radiative transfer is required.

Box 1.2. The radiative energy density is $U = aT^4$ [PHAS 2112, eqtn. (3.27)], so that $\mathrm{d}U/\mathrm{d}T = 4aT^3$, and we can express eqtn. (1.13) as

$$\begin{aligned}F &= \int_0^\infty F_\nu \, \mathrm{d}\nu \\ &= -\frac{c}{3\bar{k}_\mathrm{R}}\frac{\mathrm{d}T}{\mathrm{d}r}\frac{\mathrm{d}U}{\mathrm{d}T} \\ &= -\frac{c}{3\bar{k}_\mathrm{R}}\frac{\mathrm{d}U}{\mathrm{d}r}\end{aligned}$$

This 'diffusion approximation' shows explicitly how the radiative flux relates to the energy gradient; the constant of proportionality, $c/3\bar{k}_\mathrm{R}$, is called the diffusion coefficient. The larger the opacity, the less the flux of radiative energy, as one might intuitively expect.

1.3.2 Radiative temperature gradient

The stellar luminosity at some radius r is given by

$$L(r) = 4\pi r^2 \int_0^\infty F_\nu \, \mathrm{d}\nu$$

so, finally,

$$L(r) = -\frac{16\pi}{3} \frac{r^2}{\overline{k}_\mathrm{R}} \frac{\mathrm{d}T}{\mathrm{d}r} acT^3, \tag{1.15}$$

We can simply rearrange eqtn. (1.15) to express the temperature gradient where energy transport is radiative:

$$\frac{\mathrm{d}T}{\mathrm{d}r} = -\frac{3}{16\pi} \frac{\overline{k}_\mathrm{R}}{r^2} \frac{L(r)}{acT^3} = -\frac{3}{16\pi} \frac{\overline{\kappa}_\mathrm{R}\rho(r)}{r^2} \frac{L(r)}{acT^3}. \tag{1.16}$$

Combining this result with hydrostatic equilibrium,

$$\frac{\mathrm{d}P(r)}{\mathrm{d}r} = \frac{-Gm(r)\rho(r)}{r^2}, \tag{1.1}$$

we obtain

$$\frac{\mathrm{d}T}{\mathrm{d}P} = -\frac{3\overline{\kappa}_\mathrm{R} L(r)}{16\pi \, acT^3 \, Gm(r)} \tag{1.17}$$

or equivalently, in a form that will be of use later,

$$\frac{\mathrm{d}\ln T}{\mathrm{d}\ln P} = -\frac{3\overline{\kappa}_\mathrm{R} L(r) P}{16\pi \, acT^4 \, Gm(r)} \tag{1.18}$$

1.3.3 Von Zeipel's law

From eqtn. (1.13),

$$F \propto \frac{T^3}{\overline{\kappa}_\mathrm{R}\,\rho} \frac{\mathrm{d}T}{\mathrm{d}r} \tag{1.19}$$

$$\propto \frac{T^3}{\overline{\kappa}_\mathrm{R}\,\rho} \frac{\mathrm{d}T}{\mathrm{d}\psi} \frac{\mathrm{d}\psi}{\mathrm{d}r} \tag{1.20}$$

where ψ is the gravitational potential (and hence $\mathrm{d}\psi/\mathrm{d}r$ is the local gravity,[4] g). In hydrostatic equilibrium (see eqtn. 1.1)

$$\frac{\mathrm{d}P}{\mathrm{d}r} = -\rho(r)g(r) \propto \rho\frac{\mathrm{d}\psi}{\mathrm{d}r} \tag{1.1}$$

[4] In circumstances where von Zeipel's law is important, gravity is, in general, not a central force, so we should actually set $g = \nabla\psi$; but the central-force approximation is adequate for our purposes (and the correct general result is obtained)

so that the pressure P is a function of the potential ψ – and hence the density must also be a function of ψ.[5] For an equation of state of the general form

$$T = T(P, \rho) \tag{1.21}$$

we therefore see that T must also be a function of ψ,

$$T = T(\psi). \tag{1.22}$$

The coeffiecent of $\mathrm{d}\psi/\mathrm{d}r$ in eqtn. (1.20) is therefore a function of ψ alone, whence

$$F \propto \frac{\mathrm{d}\psi}{\mathrm{d}r} \propto g \tag{1.23}$$

or, equivalently,

$$T_{\text{eff}} \propto g^{0.25} \tag{1.24}$$

which is known as von Zeipel's law. Although it relies on the assumption of radiative energy transport by diffusion, which breaks down in a stellar atmosphere, the atmosphere is usually very thin compared to the radiative envelope, so the *surface* flux can be expected to obey eqtn. (5.18) for stars in hydrostatic equilibrium and for which energy transport through the outer envelope is radiative.

Von Zeipel's law is of particular interest for close binary stars and rapidly rotating single stars. In either case, the local gravity, and hence the local temperature, can vary over the stellar surface (which is defined by a constant potential). Although increasing gravity results in increasing flux, the effect the practical effects have come to be known as gravity *darkening*, because rapid rotation, or a close companion star, both serve to reduce a star's local gravity (and hence reduce the temperature locally).

It's of interest that von Zeipel also demonstrated that a rotating star *cannot* be simultaneously in strict hydrostatic and radiative equilibrium, undermining the basis of his 'law'. In practice, as shown by Eddington and by Sweet, rotation induces circulation currents in the stellar interior; however, these currents are sufficiently slow as to not lead to significant departures from hydrostatic equilibrium (the circulation timescales are long compared to the dynamical timescales discussed in Section 14.1.1), and gravity darkening is observed to occur in practice.

1.3.4 Mass–Luminosity Relationship

We can put together our basic stellar-structure relationships to demonstrate a scaling between stellar mass and luminosity. From hydrostatic equilibrium,

$$\frac{\mathrm{d}P(r)}{\mathrm{d}r} = \frac{-Gm(r)\rho(r)}{r^2} \quad \rightarrow \quad P \propto \frac{M}{R}\rho \tag{1.1}$$

[5] Since ρ is a scalar, the gradients of P and ψ are everywhere parallel.

but our (gas) equation of state is $P = (\rho k T)/(\mu m(\mathrm{H}))$, so

$$T \propto \frac{\mu M}{R}.$$

For stars in which the dominant energy transport is radiative, we have

$$L(r) \propto \frac{r^2}{\overline{k}_\mathrm{R}} \frac{\mathrm{d}T}{\mathrm{d}r} T^3 \qquad \propto \frac{r^2}{\overline{\kappa}_\mathrm{R} \rho(r)} \frac{\mathrm{d}T}{\mathrm{d}r} T^3 \tag{1.15}$$

so that

$$L \propto \frac{R T^4}{\overline{\kappa}_\mathrm{R} \rho}.$$

From mass continuity (or by inspection) $\rho \propto M/R^3$, giving

$$\begin{aligned} L &\propto \frac{R^4 T^4}{\overline{\kappa}_\mathrm{R} M} \\ &\propto \frac{R^4}{\overline{\kappa}_\mathrm{R} M} \left(\frac{\mu M}{R} \right)^4; \end{aligned}$$

i.e.,

$$L \propto \frac{\mu^4}{\overline{\kappa}_\mathrm{R}} M^3.$$

This simple dimensional analysis yields a dependency which is in quite good agreement with observations; for solar-type main-sequence stars, the empirical mass–luminosity relationship is $L \propto M^{3.5}$.

1.4 Convection in stars

1.4.1 Schwarzschild criterion

For convection to occur, there must be some temperature gradient (in the case of stars, a radial temperature gradient). We have seen that the temperature gradient where energy transport is radiative is given by

$$\frac{\mathrm{d}T}{\mathrm{d}r} = -\frac{3}{16\pi} \frac{\overline{k}_\mathrm{R}}{r^2} \frac{L(r)}{acT^3}; \tag{1.16}$$

that is, high opacity leads to large temperature gradients (as we might expect intuitively; the opacity block sthe flow of radiant energy from hotter to coller regions). If the energy flux isn't contained by the temperature gradient, we have to invoke another mechanism – convection – for energy transport (recall, conduction is negligible in ordinary stars.) Under what circumstances will this arise? Karl Schwarzschild (1906) developed a standard criterion for

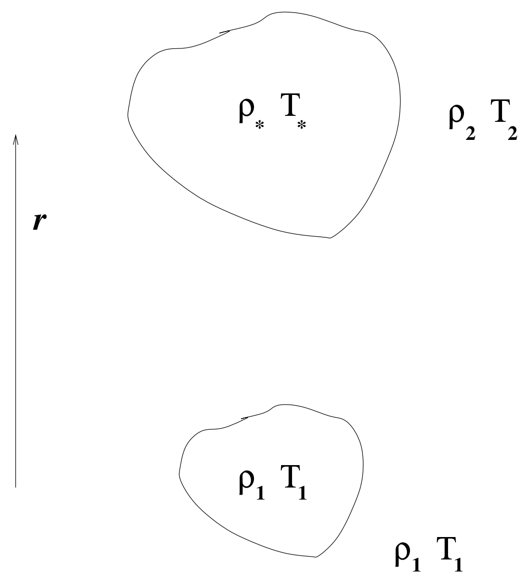

Figure 1.2: A (potentially) convective 'blob' in a stellar envelope, defining the terminology used in Section 1.4.

determining if convection occurs or not. (Here we'll derive it as a criterion for stability, although we could equally well establish a criterion for instability.)

To follow Schwarzschild's reasoning, we suppose that we start with a stellar envelope in radiative equilibrium – in some sense, its 'natural state' – and that, through some minor perturbation, an element (or cell, or blob, or bubble) of gas is displaced upwards within a star. Our essential assumptions are that the cell rises subsonically, so that hydrostatic equilibrium (or, equivalently, pressure equilibrium) is maintained; and that the cell cools adiabatically (but that the ambient temperature is determined by radiative equilibrium).

As the cell rises into a lower-pressure regime, it will expand to bring it into pressure equilibrium with the surroundings (a process whose timescale is naturally set by the speed of sound and the linear scale of the perturbation), but not, in general, into thermal equilibrium; that is, its pressure, but not its density and temperature, will match conditions in the surrounding gas. If it cell gas is less dense, then simple buoyancy comes into play; the cell will continue to rise, and convective motion occurs.[6]

We obtain *stability* (rising cell denser than surroundings[7]) if

$$|\Delta\rho_{\rm ad}| < |\Delta\rho_{\rm rad}|$$

[6]Another way of looking at this is that the entropy (per unit mass) of the blob is conserved, so the star is unstable if the ambient entropy per unit mass decreases outwards.

[7]We could follow identical arguments for stability by requiring a descending cell to be less dense than its surroundings

(where the 'ad', 'rad' subscripts indicate adiabatic and radiative conditions), that is, if

$$\left|\frac{\mathrm{d}\rho}{\mathrm{d}r}\right|_{\mathrm{ad}} < \left|\frac{\mathrm{d}\rho}{\mathrm{d}r}\right|_{\mathrm{rad}} \tag{1.25}$$

since

$$\Delta\rho = \left(\frac{\mathrm{d}\rho}{\mathrm{d}r}\right)\Delta r$$

(and Δr is the same for the cell and the ambient gas).

We now use our assumption of pressure equilibrium to express this result in terms of temperature (instead of density); the change in pressure between r_1 and r_2 is the same inside the cell as outside, so

$$\Delta P_{\mathrm{ad}} = \Delta P_{\mathrm{rad}};$$

but $P \propto \rho T$ (equation of state, eqtn. 1.2), so

$$\Delta\rho_{\mathrm{ad}} T_{\mathrm{ad}} = \Delta\rho_{\mathrm{rad}} T_{\mathrm{rad}}.$$

In other words, an increase in density is matched by a decrease in temperature, hence

$$\left|\frac{\mathrm{d}T}{\mathrm{d}r}\right|_{\mathrm{ad}} > \left|\frac{\mathrm{d}T}{\mathrm{d}r}\right|_{\mathrm{rad}} \tag{1.26}$$

is equivalent to eqtn. (1.25) – i.e., is the condition for stability.

Finally, we invoke the equation of hydrostatic equlibrium

$$\frac{\mathrm{d}P(r)}{\mathrm{d}r} = -\rho(r)\, g(r) \tag{1.1}$$

and the (gas-pressure) equation of state,

$$P = (\rho k T)/(\mu m(\mathrm{H})) \tag{1.2}$$

to write

$$\begin{aligned}
\left|\frac{\mathrm{d}T}{\mathrm{d}r}\right| &= \left|\frac{\mathrm{d}T}{\mathrm{d}P}\frac{\mathrm{d}P}{\mathrm{d}r}\right| \\
&= \left|\frac{\mathrm{d}T}{\mathrm{d}P}\; g\rho\right| \\
&= \left|\frac{\mathrm{d}T}{\mathrm{d}P}\right|\; g\frac{\mu m(\mathrm{H})}{kT}P \\
&= \left|\frac{\mathrm{d}\ln T}{\mathrm{d}\ln P}\right|\; g\frac{\mu m(\mathrm{H})}{k} \qquad (1.27)
\end{aligned}$$

Substituting this into eqtn. (1.26) we obtain

$$\left|\frac{\mathrm{d}\ln T}{\mathrm{d}\ln P}\right|_{\mathrm{ad}} > \left|\frac{\mathrm{d}\ln T}{\mathrm{d}\ln P}\right|_{\mathrm{rad}}$$

which is frequently written in the more compact notation

$$\nabla_{\mathrm{ad}} > \nabla_{\mathrm{rad}} \tag{1.28}$$

which is the *Schwarzschild criterion* for stability. It tells us that if the temperature gradient in the stellar envelope is larger than the adiabatic temperature gradient, convection occurs.

1.4.2 What does this mean in practice?

Since large temperature gradients arise in (initially) radiative envelopes if the opacity is high (eqtn. 1.16), we interpret this as meaning that convection occurs when the opacity is too high for radiative transport to be efficient; but how do we evaluate $|\mathrm{d}(\ln T)/\mathrm{d}(\ln P)|$? We appeal to thermodynamics.

Under adiabatic conditions

$$PV^{\gamma} = \text{constant}$$

where $\gamma = C_P/C_V$, the ratio of specific heats at constant pressure and constant volume. Thus, for a gas cell of constant mass ($V \propto \rho^{-1}$),

$$\begin{aligned} P &\propto \rho^{\gamma}; && \text{but also} \\ P &\propto \rho T, && \text{so that} \\ P^{\gamma-1} &\propto T^{\gamma} && \end{aligned} \tag{1.2}$$

and so

$$\left|\frac{\mathrm{d}(\ln T)}{\mathrm{d}(\ln P)}\right|_{\mathrm{ad}} = \frac{\gamma-1}{\gamma} \tag{1.29}$$

The Schwarzschild criterion for stability can therefore be written as

$$\begin{aligned} &\left|\frac{\mathrm{d}(\ln T)}{\mathrm{d}(\ln P)}\right|_{\mathrm{rad}} < \left|\frac{\mathrm{d}(\ln T)}{\mathrm{d}(\ln P)}\right|_{\mathrm{ad}} = \frac{\gamma-1}{\gamma} \\ \equiv\quad &\nabla_{\mathrm{rad}} \qquad\quad < \quad \nabla_{\mathrm{ad}} < \frac{\gamma-1}{\gamma} \end{aligned} \tag{1.30}$$

or, in terms of temperature gradient (cp. eqtn. 1.27)

$$\left|\frac{\mathrm{d}T}{\mathrm{d}r}\right|_{\mathrm{rad}} < \left|\frac{\mathrm{d}T}{\mathrm{d}r}\right|_{\mathrm{ad}} = \frac{\gamma-1}{\gamma}\left|\frac{T}{P}\frac{\mathrm{d}P}{\mathrm{d}r}\right|. \tag{1.31}$$

We know the radiative temperature gradient (eqtn. 1.16); whence, by reference to eqtns (1.18) and (1.29). the Schwarzschild criterion for stability can be written as

$$\frac{3\overline{\kappa}_{\rm R}L(r)P}{16\pi acT^4Gm(r)} < \frac{\gamma-1}{\gamma} \tag{1.32}$$

Finally, yet another version can be obtained by noting that the equation of state $P = nkT = (\rho kT)/(\mu m(\mathrm{H}))$ gives

$$\frac{\mathrm{d}\ln\rho}{\mathrm{d}\ln P} = 1 + \frac{\mathrm{d}\ln\mu}{\mathrm{d}\ln P} - \frac{\mathrm{d}\ln T}{\mathrm{d}\ln P} \tag{1.33}$$

whence the Schwarzschild criterion for stability is

$$\frac{\mathrm{d}\ln T}{\mathrm{d}\ln P}_{\rm ad} < \frac{\gamma-1}{\gamma} + \frac{\mathrm{d}\ln\mu}{\mathrm{d}\ln P} \tag{1.34}$$

(demonstrating explicitly that compositional changes can influence whether or not convenction occurs).

1.4.3 Physical conditions for convection

From equations 1.30–1.34 we can see several ways in which convection may, in principle be induced, but eqtn. (1.30) argues that the essential requirements are either:
$\nabla_{\rm rad}$ becomes large (compared to $\nabla_{\rm ad}$), or
$\nabla_{\rm ad}$ becomes small (compared to $\nabla_{\rm rad}$). Alternatively, from eqtn. (1.32), we can see what this means in terms of luminosity, opacity, and the adiabatic exponent γ.

In nature, convectively unstable regions occur:

(i) In the cores of massive stars, where the radiation flux $L(r)/4\pi r^2$ can be very large, driving convection. The opacity is too great to allow the radiation to flow at an equilibrium rate (the κ effect).

(ii) In the outer regions of cool stars, where the adiabatic exponent γ can approach unity (and hence $(\gamma-1)/\gamma$ can become very small; the γ effect).

For a monatomic ideal gas (representative of stellar interiors),[8] $\gamma = 5/3$ and so $(\mathrm{d}\ln T/\mathrm{d}\ln P)_{\rm ad} = 0.4$, but under changing conditions of ionization this exponent changes. For a simple pure-hydrogen composition it can be shown that

$$\left|\frac{\mathrm{d}\ln T}{\mathrm{d}\ln P}\right| = \frac{2 + X(1-X)\left((5/2) + E_1/(kT)\right)}{5 + X(1-X)\left((5/2) + E_1/(kT)\right)^2} \tag{1.35}$$

[8] Radiation obeys a 'gas law' with $\gamma = 4/3$

where

$$X = \frac{n_{\rm e}}{n_{\rm P} + n({\rm H}^0)} \tag{1.36}$$

is the degree of ionization, and E_1 is the ionization potential. For $X = 0$ or 1, this recovers $(\mathrm{d}\ln T/\mathrm{d}\ln P)_{\rm ad} = 0.4$, but in regions of partial ionization lower values apply, with a minimum at $X = 0.5$ [$(\mathrm{d}\ln T/\mathrm{d}\ln P)_{\rm ad} = 0.07$] which occurs (e.g.) near the base of the solar photosphere.

The switch from radiative core/convective envelope to convective core/radiative envelope occurs on the main sequence at masses only very slightly more than the Sun's. This is related to the core energy-generation mechanism, as the principal hydrogen-burning process switches from proton-proton chains (which generate energy at a rate that can be transported radiatively) to CNO processing.

1.4.4 Convective energy transport: mixing-length 'theory'

So far, we have only tested whether or not convection is likely to occur; we have not addressed the convective flux. Unfortunately, convection is a complex, hydrodynamic process. Although much progress is being made in numerical modelling of convection over short timescales, it's not feasible at present to model convection in detail in stellar-evolution codes routinely, because of the vast disparities between convective and evolutionary timescales. Instead, we appeal to simple parameterizations of convection, of which mixing-length 'theory' is the most venerable, and the most widely applied.

We again suppose that the envelope becomes convectively unstable at some radius r_0, and that the cell then rises through some characteristic distance ℓ - the *mixing length*;
the excess thermal energy of the cell is released into the ambient medium; and
the cooled cell sinks back down.
Because we are moving energy from deeper to shallower regions, the temperature gradient is shallower for the cell than the pure radiative case.

From hydrostatic equilibrium (eqtn. 1.1) and the perfect gas equation (eqtn. 1.2) we have

$$\begin{aligned}\frac{\mathrm{d}P}{\mathrm{d}r} &= -gP\frac{\mu m(\mathrm{H})}{kT}, \qquad \text{or}\\ \frac{\mathrm{d}P}{P} &= -g\frac{\mu m(\mathrm{H})}{kT}\mathrm{d}r, \equiv -\frac{\mathrm{d}r}{H}.\end{aligned} \tag{1.37}$$

The solution of eqtn. 1.37 is

$$P = P_0 \exp\left(-r/H\right)$$

so H, the pressure scale height, is the vertical distance over which the pressure drops by a factor e. The mixing length is conveniently expressed in terms of this scale height; typically, we expect $\ell \simeq H$, but since the detailed physics is not well understood ascale factor is usually introduced, whereby

$$\ell = \alpha H,$$

with $\alpha \sim 0.5$–1.5.

For simplicity (justified given the weakness of other assumptions), we suppose that ℓ is the same for all cells, and that the velocity of all cells is also the same.

Now, the excess temperature of the cell (compared to the ambient gas) is, in general,

$$\Delta T = \left\{ \left| \frac{\mathrm{d}T}{\mathrm{d}r} \right|_{\mathrm{rad}} - \left| \frac{\mathrm{d}T}{\mathrm{d}r} \right|_{\mathrm{ad}} \right\} \times \Delta r \tag{1.38}$$

For a cell moving with velocity v the flux of energy across unit area is given by the mass flux times the heat energy per unit mass:

$$\begin{aligned} \pi F_{\mathrm{conv}} &= \rho v \times \mathrm{d}Q \\ &= \rho v \times C_{\mathrm{P}} \Delta T \end{aligned} \tag{1.39}$$

where C_{P} is the specific heat at constant pressure; so we need an estimate of v. We obtain this by considering the buoyancy force,

$$f_{\mathrm{b}} = -g \Delta\rho \tag{1.40}$$

where $\Delta\rho$ is the density difference between the cell and ambient gas. Then from the equation of state, eqtn. 1.2,

$$\frac{\Delta P}{P} = \frac{\Delta\rho}{\rho} + \frac{\Delta T}{T} - \frac{\Delta\mu}{\mu} \tag{1.41}$$

but in pressure equlibrium $\Delta P = 0$, whence

$$\frac{\Delta\rho}{\rho} = \frac{\Delta\mu}{\mu} - \frac{\Delta T}{T}$$

or (taking a limit)

$$\Delta\rho = -\rho \frac{\Delta T}{T} \left(1 - \frac{\mathrm{d}\ln\mu}{\mathrm{d}\ln T} \right) \tag{1.42}$$

so the buoyancy force, eqtn. 1.40, is

$$f_{\mathrm{b}} = g\rho \frac{\Delta T}{T} \left(1 - \frac{\mathrm{d}\ln\mu}{\mathrm{d}\ln T} \right)$$

but force equals mass (per unit volume) times acceleration,

$$= \rho \frac{\mathrm{d}v}{\mathrm{d}t} \tag{1.43}$$

so

$$\frac{\mathrm{d}v}{\mathrm{d}t} = g\frac{\Delta T}{T}\left(1 - \frac{\mathrm{d}\ln\mu}{\mathrm{d}\ln T}\right) \tag{1.44}$$

For constant acceleration,

$$v \simeq \sqrt{\left\{\frac{\mathrm{d}v}{\mathrm{d}t}\ell\right\}} \tag{1.45}$$

so, substituting eqtn. 1.38 for ΔT in eqtn. 1.44 (setting $\Delta r = \ell$), the required velocity is

$$v = \left\{\frac{g}{T}\left|1 - \frac{\mathrm{d}\ln\mu}{\mathrm{d}\ln T}\right|\right\}^{1/2} \left\{\left|\frac{\mathrm{d}T}{\mathrm{d}r}\right|_{\mathrm{rad}} - \left|\frac{\mathrm{d}T}{\mathrm{d}r}\right|_{\mathrm{ad}}\right\}^{1/2} \times \ell \tag{1.46}$$

We can now rewrite eqtn. 1.39 as

$$\pi F_{\mathrm{conv}} = \rho\, C_{\mathrm{P}} \left\{\frac{g}{T}\left|1 - \frac{\mathrm{d}\ln\mu}{\mathrm{d}\ln T}\right|\right\}^{1/2} \left\{\left|\frac{\mathrm{d}T}{\mathrm{d}r}\right|_{\mathrm{rad}} - \left|\frac{\mathrm{d}T}{\mathrm{d}r}\right|_{\mathrm{ad}}\right\}^{3/2} \times \ell^2. \tag{1.47}$$

Rearranging the equation of state, eqtn. (1.2),

$$\begin{aligned} \left|\frac{\mathrm{d}T}{\mathrm{d}r}\right| &= \frac{g\mu m(\mathrm{H})}{k}\left|\frac{\mathrm{d}\ln T}{\mathrm{d}\ln P}\right| \\ &= \frac{T}{H}\left|\frac{\mathrm{d}\ln T}{\mathrm{d}\ln P}\right| \end{aligned} \tag{1.48}$$

and

$$\pi F_{\mathrm{conv}} = \rho\, C_{\mathrm{P}}\, \alpha^2 T \left\{gH\left|1 - \frac{\mathrm{d}\ln\mu}{\mathrm{d}\ln T}\right|\right\}^{1/2} \left\{\left|\frac{\mathrm{d}\ln T}{\mathrm{d}\ln P}\right|_{\mathrm{rad}} - \left|\frac{\mathrm{d}\ln T}{\mathrm{d}\ln P}\right|_{\mathrm{ad}}\right\}^{3/2} \tag{1.49}$$

In calculating actual temperature structures in stellar envelopes, we require the total energy flux to obey

$$\pi F = \pi F_{\mathrm{rad}} + \pi F_{\mathrm{conv}} = \sigma T_{\mathrm{eff}}^4 \tag{1.50}$$

The initial temperature structure is calculated on the basis of radiative transfer only ($\pi F_{\mathrm{rad}} = \sigma T_{\mathrm{eff}}^4$), then a correction $\Delta T(r)$ computed iteratively, for given α, if the Schwarzschild criterion indicates convective transport.