Chapter 8
The Chemical Composition

8.1 Relative Mass Abundances

The chemical composition of stellar matter is obviously very important, since it directly influences such basic properties as absorption of radiation or generation of energy by nuclear reactions. These reactions in turn alter the chemical composition, which represents a long-lasting record of the nuclear history of the star.

The composition of stellar matter is extremely simple compared to that of terrestrial bodies. Because of the high temperatures and pressures, there are no chemical compounds in the stellar interior, and the atoms are for the most part completely ionized. It suffices then to count and keep track of the different types of nuclei.

We denote by X_i that fraction of a unit mass which consists of nuclei of type i. This requires that

$$\sum_i X_i = 1 \, . \tag{8.1}$$

The chemical composition of a star at time t is then described, if for the relevant nuclei the functions $X_i = X_i(m, t)$ are given in the interval $[0, M]$ of m.

The commonly used particle number per volume, n_i, of nuclei with mass m_i, is related to the mass abundances by

$$X_i = \frac{m_i n_i}{\varrho} \, . \tag{8.2}$$

Usually, one does not need to specify very many X_i because most elements are either too rare or play no relevant role, or their abundances remain constant in time. In fact, for many purposes, it is even sufficient to specify only the mass fractions of hydrogen, helium, and "the rest" with the notation

$$X \equiv X_{\mathrm{H}}, \quad Y \equiv X_{\mathrm{He}}, \quad Z \equiv 1 - X - Y \, . \tag{8.3}$$

R. Kippenhahn et al., *Stellar Structure and Evolution*, Astronomy and Astrophysics Library, DOI 10.1007/978-3-642-30304-3_8, © Springer-Verlag Berlin Heidelberg 2012

This requires additional conventions about the relative distribution of the elements in Z, collectively called "metals", in particular the amount of C, N, and O, which are important for hydrogen burning.

Young stars throughout, and most stars in their envelopes, contain an overwhelming amount of hydrogen and helium: $X = 0.65 \ldots 0.75$, $Y = 0.30 \ldots 0.25$, $Z = 0.05 \ldots 0.0001$.

Of course, nuclear reactions will eventually change this simple picture drastically. For example, if many competing reactions occur simultaneously, or if one is interested in such aspects as isotopic ratios, one may have to specify a large number of different X_i. Only if inverse β decay, the big equalizer in late stages of evolution, has destroyed all elements does the composition then return to utmost simplicity–just neutrons (Chap. 38).

The advantages of the use of m instead of r as independent variable become particularly evident when we have to describe the chemical composition. If we took $X_i(r, t)$ instead, any expansion would immediately lead to a change of all the functions X_i, this holds, of course, for all functions depending on the chemical composition.

8.2 Variation of Composition with Time

8.2.1 Radiative Regions

In radiative regions, there is no exchange of matter between different mass shells, if we can neglect diffusion. Then, the X_i can change only if nuclear reactions create or destroy nuclei of type i in the mass element under consideration.

The frequency of a certain reaction is described by the *reaction rate* r_{lm}, that is, the number of reactions per unit volume and time that transform nuclei from type l into type m (see Chap. 18). The reaction itself will in most cases involve more than just one mother and one daughter nucleus, but for simplicity, we characterize it by one index only. In general, an element i can be affected simultaneously by many reactions, some of which create it (r_{ji}) and some of which destroy it (r_{ik}). These reaction rates give directly the change per second of n_i. Then, with (8.2), we have

$$\frac{\partial X_i}{\partial t} = \frac{m_i}{\varrho} \left[\sum_j r_{ji} - \sum_k r_{ik} \right], \quad i = 1 \ldots I \tag{8.4}$$

for any of the elements $1 \cdots I$ which are involved in reactions (If more than one nucleus of type i is created or destroyed per reaction, the corresponding terms in the sums have simply to be normalized by the number of nuclei of type i involved.).

The reaction $p \rightarrow q$ in which one nucleus of type p is transformed may be connected with a release of energy e_{pq}. In the equation of energy conservation, we have used the energy generation rate ε per unit mass, which normally contains contributions from several different reactions. The ε are simply proportional to the reaction rates:

$$\varepsilon = \sum_{p,q} \varepsilon_{pq} = \frac{1}{\varrho} \sum_{p,q} r_{pq}\, e_{pq} \ . \tag{8.5}$$

Let us introduce the energy generated when one mass unit of type p nuclei is transformed into type q:

$$q_{pq} = \frac{e_{pq}}{m_p} \ . \tag{8.6}$$

For simple cases, it is convenient to rewrite (8.4) in terms of the ε, which already occur in the equation of energy conservation. If all reactions give a positive contribution to ε, then instead of (8.4), we can write

$$\frac{\partial X_i}{\partial t} = \sum_j \frac{\varepsilon_{ji}}{q_{ji}} - \sum_k \frac{\varepsilon_{ik}}{q_{ik}} \ . \tag{8.7}$$

If I different nuclei are simultaneously subject to nuclear transformations, equations (8.4) or (8.7) form a set of I differential equations, technically called a "nuclear reactions network". One of them could be replaced by the normalization (8.1), such that we need only $I - 1$ of them to complete the basic equations of our problem. Technically, however, this is not advisable, as (8.1) can then serve as an independent consistency check: if the set of differential equations is solved correctly, mass must be conserved.

Note that for simple cases, it may even suffice to consider just one of these equations. For example, if hydrogen burning is to be taken into account only by way of an overall generation rate ε_{H} (giving the sum over all single reactions), then the only equation needed is

$$\frac{\partial X}{\partial t} = -\frac{\varepsilon_{\mathrm{H}}}{q_{\mathrm{H}}} \tag{8.8}$$

with $\partial Y/\partial t = -\partial X/\partial t$, where q_{H} is the energy release per unit mass when hydrogen is converted into helium.

In Sect. 4.6, we defined the nuclear timescale for a certain burning, $\tau_{\mathrm{n}} = E_{\mathrm{n}}/L$. One can actually define a nuclear timescale for each type of nuclear burning since each nuclear energy reservoir is proportional to an integral of $X_i \cdot dm$ over the whole star, where X_i refers to the element consumed by the reactions; therefore, τ_{n} is equivalent to τ_{Xi}, the timescale for the exhaustion of the element i.

8.2.2 *Diffusion*

Certain microscopic effects can also change the chemical composition in a star. If gradients occur in the abundances of chemical elements, then *concentration diffusion* tends to smooth out the differences. Even in chemically homogeneous stellar layers, heavier atoms can migrate towards the regions of higher temperature, owing to the effect of *temperature diffusion.* Also, the pressure gradient in a stratified layer causes the heavier particles to diffuse towards the region of higher pressure, that is, *pressure diffusion*. The detailed statistical theory of diffusion is derived in Burgers (1969), Chapman and Cowling (1970), and Choudhuri (1998).

We start with the simplest case: concentration diffusion. Let c be the concentration of particles of a certain species, that is, the number density of particles of that type divided by the number density of all particles, and $\boldsymbol{j}_{\mathrm{D}}$ be the "flux of concentration"; then, Fick's first law states that

$$\boldsymbol{j}_{\mathrm{D}} = -D\nabla c \ , \tag{8.9}$$

where D is the diffusion coefficient (We will derive (8.9) later.). With $\boldsymbol{j}_{\mathrm{D}} = c\boldsymbol{v}_{\mathrm{D}}$, where $\boldsymbol{v}_{\mathrm{D}}$ is the diffusion velocity, one has

$$\boldsymbol{v}_{\mathrm{D}} = -\frac{D}{c}\nabla c \ . \tag{8.10}$$

With the continuity equation

$$\frac{\partial c}{\partial t} = -\nabla \cdot \boldsymbol{j}_{\mathrm{D}}, \tag{8.11}$$

we find that

$$\frac{\partial c}{\partial t} = \nabla \cdot (D\nabla c) \ , \tag{8.12}$$

and in the case of constant D that

$$\frac{\partial c}{\partial t} = D\nabla^2 c \ , \tag{8.13}$$

a rough estimate for the characteristic timescale is given by

$$\tau_{\mathrm{D}} \approx \frac{S^2}{D} \ , \tag{8.14}$$

where S is a characteristic length for the variation of c.

By generalizing (8.10) one can formally include the two other types of diffusion, i.e.

$$\boldsymbol{v}_{\mathrm{D}} = -\frac{1}{c}D(\nabla c + k_T \nabla \ln T + k_P \nabla \ln P) \ , \tag{8.15}$$

if the coefficients k_T and k_P are properly specified. In order to do that we first consider the combined effects of concentration and temperature diffusion.

We assume ∇T to be perpendicular to the x–y plane in a Cartesian coordinate system; then the flux of particles of a certain type in the $+z$ direction due to the statistical motion of the particles is determined by the density n and the mean velocity $\overline{v}$, both taken at $z = -\ell$, where ℓ is the mean free path of the particles of this type:

$$j^+ = \frac{1}{6} c(-\ell)\overline{v}(-\ell) \,, \tag{8.16}$$

where the numerical factor originates in averaging over $\cos^2$. This takes into account that the particles penetrating the x–y plane had their last encounter at $z = -\ell$.

If one expands n and $\overline{v}$ at $z = 0$ in (8.16) and in a corresponding expression for j^-, the fluxes in the $+z$ and $-z$ directions are

$$j^\pm = \frac{1}{6}\left(c(0) \mp \frac{\partial c}{\partial z}\ell\right)\left(\overline{v}(0) \mp \frac{\partial \overline{v}}{\partial z}\ell\right) \,, \tag{8.17}$$

and therefore there is a net flux

$$j = j^+ - j^- = -\frac{1}{3}\left(\frac{\partial c}{\partial z}\ell\overline{v} + \frac{\partial \overline{v}}{\partial z}\ell c\right) \,, \tag{8.18}$$

which in general does not vanish, i.e. we have obtained Fick's law.

We now consider the relative diffusion velocity $v_{\mathrm{D}_1} - v_{\mathrm{D}_2}$ resulting from the motion of two different types of particles (1, 2), with fluxes j_1, j_2 and concentrations c_1, c_2:

$$v_{\mathrm{D}_1} - v_{\mathrm{D}_2} = \frac{j_1}{c_1} - \frac{j_2}{c_2} \,. \tag{8.19}$$

With (8.18) we can replace the j_i by $\ell_i, \overline{v}_i$, and the gradients of c_i, while the velocity gradient–with the help of $\overline{v}_i = (3\Re T/\mu_i)^{1/2}$–can be replaced by the temperature gradient. Using the continuity equation (and after some algebra) an expression of the form

$$v_{\mathrm{D}_1} - v_{\mathrm{D}_2} = -\frac{D}{c_1 c_2}\left(\frac{\partial c_1}{\partial z} + k_T \frac{\partial \ln T}{\partial z}\right) \tag{8.20}$$

follows. The two terms in the brackets are responsible for concentration diffusion and temperature diffusion. In a mixture of two species ($i = 1, 2$) D and k_T have the form

$$D = \frac{1}{3}(c_2\ell_1\overline{v}_1 + c_1\ell_2\overline{v}_2) = \left(\frac{\Re T}{3}\right)^{1/2}\left(c_2\ell_1\mu_1^{-1/2} + c_1\ell_2\mu_2^{-1/2}\right) \,, \tag{8.21}$$

$$k_T = \frac{1}{2}\frac{\ell_1\sqrt{\mu_2} - \ell_2\sqrt{\mu_1}}{\ell_1 c_2\sqrt{\mu_2} + \ell_2 c_1\sqrt{\mu_1}} c_1 c_2 (c_2 - c_1) \,, \tag{8.22}$$

where ℓ_1 and ℓ_2 are the mean free paths of the two species (Landau and Lifshitz 1987, vol. 6). The absolute value k_T is of order 1 or less, and its sign is not immediately clear, though more detailed considerations indicate that $k_T > 0$ for a typical ionized hydrogen–helium mixture in stars.

From (8.21) it is obvious that D is of order

$$D \approx \left(\frac{\Re T}{3}\right)^{1/2} \ell \approx \frac{1}{3} v^* \ell \,, \tag{8.23}$$

where v^* and ℓ are some kind of averages of the statistical velocities and the mean free paths of both components. This expression for D can be used to estimate the timescale τ_D according to (8.14). As long as $|k_T| \approx 1$ this also gives the characteristic timescale for temperature diffusion.

Since $D > 0$, in the case of $k_T > 0$ for pure temperature diffusion, one has $\mathrm{sign}(v_\mathrm{D}) = -\mathrm{sign}(\partial \ln T/\partial x)$. Let us now consider the case of a mixture of hydrogen and helium. Here $v_\mathrm{D} = v_\mathrm{H} - v_\mathrm{He}$ is the z component of the diffusion velocity and $v_\mathrm{D} > 0$ means that hydrogen diffuses in the direction of lower temperature, i.e. "upwards" in the star. For the central region of the Sun ($T \approx 10^7$ K, $\varrho \approx 100\,\mathrm{g\,cm^{-3}}$) one finds that $\ell \approx 10^{-8}$ cm and $D \approx 6\,\mathrm{cm^2\,s^{-1}}$, and with a characteristic length-scale $S \approx R_\odot \approx 10^{11}$ cm, the characteristic timescale τ_D (according to (8.14)) there becomes $\tau_\mathrm{D} \approx 10^{13}$ years. Although τ_D is much larger than the age of the universe and therefore the effects of concentration and temperature diffusion seem to be astrophysically irrelevant for the Sun, diffusion does have enough influence on stellar evolution such that high-precision observations require models that include its effect. This will become evident in the case of the standard solar model (see Chap. 29). We will therefore briefly discuss the situation. If a layer is homogeneous, then there is no concentration diffusion, but the hydrogen particles diffuse towards the regions of lower temperature. This causes an outward increase of n_H which in turn triggers concentration diffusion acting against the temperature diffusion ($\mathrm{sign}(\partial c_\mathrm{H}/\partial z) = -\mathrm{sign}(\partial T/\partial z)$) until both types of diffusion compensate each other.

We now turn to pressure diffusion, which is the cause of what is often called "sedimentation" or "gravitational settling". A statistical consideration similar to that used to make temperature diffusion plausible also shows that there is diffusion in isothermal layers with a non-vanishing pressure gradient. The reader is again referred to Chapman and Cowling (1970), or any of the other standard textbooks. In a way similar to that for k_T an expression for k_P in (8.15) can also be obtained.

We here confine ourselves to the discussion of the final outcome of this process of pressure diffusion, i.e. the state of final equilibrium for an isothermal layer in hydrostatic equilibrium in a gravitational field pointing towards the $-z$ direction. Let us assume that the material consists of two components ($i = 1, 2$) of perfect gases of different molecular weights μ_i and partial pressures P_i. Then there exist two pressure-scale heights $H_{P_i} = -dz/d \ln P_i$ with which (6.8) can be written in the form

$$H_{P_i} = \frac{P_i}{g\varrho_i} = \frac{\Re T}{g\mu_i} \,, \tag{8.24}$$

where $dP_i/dz = -g\varrho_i$ and $P_i = \Re\varrho_i T/\mu_i$ are used. The particle densities are proportional to the P_i, which are here approximately proportional to $\exp(-z/H_{Pi})$. Therefore the component with the higher μ_i falls off more sharply in the z direction than that with smaller μ_i, so that in a very simplified way, one can say that the heavier component has "moved below" the lighter one. This is the final state, which would be brought about by pressure diffusion alone even if the species were originally in a completely mixed state. Of course, in reality, the two other types of diffusion would also act and therefore influence the final state.

Estimates show that not only $|k_T|$ but also $|k_P|$ is of order one. Therefore it normally takes rather a long time before an appreciable separation occurs in stars. Although in general we will ignore the effect of diffusion in this book, it can be very relevant in certain special cases. Equation (8.12), using (8.15), can be formulated in terms of relative mass fractions X_i instead of particle concentrations and for the case of spherical symmetry as

$$\frac{\partial X_i}{\partial t} = -\frac{1}{\rho r^2}\frac{\partial}{\partial r}\left[r^2 X_i T^{5/2}\left(A_P(i)\frac{\partial \ln p}{\partial r} + A_T(i)\frac{\partial \ln T}{\partial r} + \sum_{k\neq e,\mathrm{He}^4} A_k(i)\frac{\partial \ln C_i}{\partial r}\right)\right]. \tag{8.25}$$

This formulation follows the one by Thoul et al. (1994), and the $T^{5/2}$ factor results from a convenient definition of the diffusion constants called A_P, A_T, and A_k here. In this description the concentration C_i is defined as c_i/c_e, i.e. as the usual particle concentration in units of the electron concentration. Note that the concentration diffusion is taken as a sum over all species, since the concentration of species i may also change due to the diffusion of all other elements. The sum actually has not to be taken over all species as mass and charge conservation reduce the number of independent A_k by two. Here we have taken out helium and electrons.

When diffusion is to be taken into account, proper evaluation of the diffusive constants D (or A in (8.25)) for the various types of diffusion is necessary. This involves correct treatment of the interaction forces between the particles and will be quite sophisticated. Two widely used sources for calculating the diffusion constants are Paquette et al. (1986) and Thoul et al. (1994), both using a method described in the book by Burgers (1969). This method is also sketched in Weiss et al. (2004). An improvement by applying quantum corrections was introduced by Schlattl and Salaris (2003). In general these diffusive speeds or constants are considered to be accurate to 20 %. There is an additional effect not discussed here: Coupling of the radiation field to partially ionized atoms results in a net upward force, counteracting the downward sedimentation. This sort of diffusion is called radiative levitation and can lead to strong variations in surface element abundances, in particular for those elements with rich energy level systems. A derivation of the relevant coefficients was given by Richer et al. (1998).

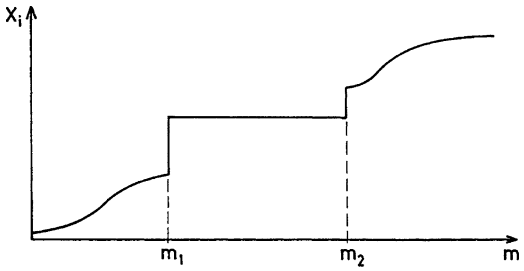

Fig. 8.1 The abundances X_i are smeared out owing to rapid mixing inside a convection zone extending from m_1 to m_2. At these borders, X_i can be discontinuous

8.2.3 Convective Regions

Here we deal with the much more important effect of mixing due to turbulent convective motion, a process that normally is very rapid compared to the extremely slow change of the chemical composition produced by nuclear reactions. Therefore we can assume that the composition in a convective region in most cases remains homogeneous,

$$\frac{\partial X_i}{\partial m} = 0 \; . \tag{8.26}$$

This requires a dispersion not only of the newly created nuclei, but of all elements inside a convective zone.

Suppose a convective zone extends between the mass values m_1 and m_2 (Fig. 8.1). Inside that interval all $X_i = \overline{X}_i$ are constant. At the boundaries one can generally have a discontinuity, such that the "outer" values X_{i1} and X_{i2} are different from the "inner" values–which are simply $\overline{X}_i$. But m_1 and m_2 can change in time, and hence one can easily see that the abundances in the convective zone vary with the rate

$$\frac{\partial \overline{X}_i}{\partial t} = \frac{1}{m_2 - m_1} \times \left(\int_{m_1}^{m_2} \frac{\partial X_i}{\partial t} dm + \frac{\partial m_2}{\partial t}(X_{i2} - \overline{X}_i) - \frac{\partial m_1}{\partial t}(X_{i1} - \overline{X}_i) \right) . \tag{8.27}$$

The X_{i1}, X_{i2} should here be taken as the value on the side that the corresponding boundary moves towards. The integral in the bracket describes the change due to nuclear reactions and can be replaced by an integral over the rates $-\varepsilon_i/q_i$, as in (8.8), where q_i is the energy released if a mass unit of the nucleus i is transformed. Without any nuclear reaction ($\partial X_i/\partial t = 0$) in the convective zone, its composition can still change if the boundaries move into a region of inhomogeneous composition, and this can have important consequences. For example, "ashes" of earlier nuclear burnings may be brought to the surface, fresh fuel may be carried into a zone of nuclear burning, or discontinuities can be produced that drastically influence the later evolution.

In cases of very fast nuclear reactions (or short nuclear timescales) the assumptions of instantaneous mixing is no longer correct. In such situations one normally treats convective mixing as a diffusive process with the diffusive velocity v_c estimated from the solution of the mixing length theory and using (7.6) and (7.16). In this case (8.25) can simply be extended by adding the additional term

$$D_c \frac{\partial X_i}{\partial r} = \left(\frac{1}{3} v_c \alpha_{\mathrm{MLT}} H_P \varrho r^2 \right) \frac{\partial X_i}{\partial r}, \tag{8.28}$$

where we used the estimate for D_c by Langer et al. (1985). Since usually D_c is by orders of magnitude larger than any of the diffusion constants in (8.25), the types of diffusion discussed in Sect. 8.2.2 can in fact be neglected in convective regions.