

# Stellar Atmospheres Theory: An Introduction

I. Hubeny

Universities Space Research Association, NASA Goddard Space Flight Center,  
Code 681, Greenbelt, MD 20771, USA

## 1 Fundamental Concepts

### 1.1 What is a Stellar Atmosphere, and Why Do We Study It?

By the term stellar atmosphere we understand any medium connected physically to a star from which the photons escape to the surrounding space. In other words, it is a region where the radiation, observable by a distant observer, originates. Since in the vast majority of cases the radiation is the only information about a distant astronomical object we have (exceptions being a direct detection of solar wind particles, neutrinos from the Sun and SN 1987a, or gravitational waves), all the information we gather about stars is derived from analysis of their radiation.

It is therefore of considerable importance to develop reliable methods which are able to decode the information about a star contained in its spectrum with confidence. Having understood the physics of the problem and being able to carry out detailed numerical simulations will enable us to construct theoretical models of a stellar atmosphere and predict a stellar spectrum. This has important applications in other branches of astrophysics, such as i) derived stellar parameters can be used to verify predictions of the stellar evolution theory; ii) models provide ionizing fluxes for the interstellar medium and nebular models; iii) predicted stellar spectra are basic blocks for population syntheses of stellar clusters, starburst regions, and whole galaxies. Moreover, very hot and massive stars have special significance. They are very bright, and therefore may be studied spectroscopically as individual objects in distant galaxies. Reliable model atmospheres for these stars may therefore yield invaluable independent information about distant galaxies, like chemical composition, and, possibly, reliable distances.

This alone would easily substantiate viewing the stellar atmosphere theory as an independent, and very important, branch of modern astrophysics. Yet, in the global astrophysical context, there is another, and equally important, contribution of the stellar atmospheres theory. Stellar atmospheres are the best studied example of a medium where radiation is not only a *probe* of the physical state, but is in fact an *important constituent*. In other words, radiation in fact *determines* the structure of the medium, yet the medium is probed *only* by this radiation.

Unlike laboratory physics, where one can change a setup of the experiment in order to examine various aspects of the studied structures separately, we do not have this luxury in astrophysics: we are stuck with the observed spectrum so we should better make a good use of it. This is exactly what the stellar atmosphere theory is doing for almost a century now. Consequently, it is developed to such an extent that it provides an excellent methodological guide for other situations where the radiation has the dual role of a probe and a constituent. Examples of such astronomical objects are the interstellar medium, H II regions, and, in particular, accretion disks.

There has been a significant progress in the field of stellar atmospheres achieved in recent years. The progress was motivated by an unprecedented increase of quality of ground- and space-based observations, and by development of extremely fast and efficient numerical methods. However, despite of this progress, the stellar atmospheres theory is still far from being sufficiently developed. It is a mature field, yet it is now reaching qualitatively new levels of sophistication. In short, it is a field worth pursuing, offering as a reward a significant contribution to our knowledge about the Universe.

The main goal of this lecture is to provide a gentle introduction to the basic concepts needed to understand the fundamental physics of stellar atmospheres, as well as the leading principles behind recent developments. Particular emphasis will be devoted to the classical plane-parallel atmospheres in hydrostatic and radiative equilibrium. Topics which concentrate specifically on non-static phenomena (stellar winds), and on departures from radiative equilibrium (stellar chromospheres and coronae), are covered in other lectures of this volume.

There is no textbook that would fully cover the topics discussed in this lecture. The fundamental textbook of the field, Mihalas (1978), is still a highly recommended text, although it does not cover important recent developments, like for instance the ALI method. The third edition of the book is now in preparation, but it will take a couple of years before it is available. There is a recent textbook by Rutten (1995), distributed electronically, which covers both the basic concepts as well as some of the modern development, and is highly recommended to the beginner in the field. There are two books edited by Kalkofen which present a collection of reviews on various mathematical and numerical aspects of radiative transfer (Kalkofen 1984; 1987). A good textbook that covers both the theoretical and observational aspects of the stellar atmospheres is that by Gray (1992). Other related textbooks include Rybicki and Lightman (1987), Shu (1991), and an elementary-level textbook by Böhm-Vitense (1989). An old but excellent textbook on radiative transfer is Jefferies (1968). Besides these books, there are several excellent review papers covering various topics (e.g. Kudritzki 1988; Kudritzki and Hummer 1990), and several conference proceedings which contain many interesting papers on the stellar atmospheres theory – Properties of Hot Luminous Stars (Garmany 1990); Stellar Atmospheres: Beyond Classical Models (Crivellari,

Hubeny, and Hummer 1991); The Atmospheres of Early-Type Stars (Heber and Jeffery 1992); and Hydrogen-Deficient Stars (Jeffery and Heber 1996); to name just few of the most important ones.

## 1.2 Basic Structural Equations

A stellar atmosphere is generally a plasma composed of many kinds of particles, namely atoms, ions, free electrons, molecules, or even dust grains, and photons. Typical values of temperature range from  $10^3$  K (or even less in the coolest stars) to a few times  $10^5$  K in the hottest stars (temperature is even higher,  $10^6 - 10^7$  K, in stellar coronae). Likewise, the total particle density ranges from, say,  $10^6$  to  $10^{16}$   $\text{cm}^{-3}$ . Under such conditions, the natural starting point for the physical description is the kinetic theory.

We start with very general equations, in order to emphasize a close connection of the stellar atmospheres theory and other branches of physics. We will then simplify these equations to the form which is used in most textbooks.

Specifically, the most general quantity which describes the system is the distribution function  $f_i(\mathbf{r}, \mathbf{p}, t)$ , which has the meaning that  $f_i(\mathbf{r}, \mathbf{p}, t) d\mathbf{r} d\mathbf{p}$  is the number of particles of kind  $i$  in an elementary volume of the phase space at position  $\mathbf{r}$ , momentum  $\mathbf{p}$ , and at time  $t$ . The equation which describes a development of the distribution function is the well-known kinetic, or Boltzmann, equation, written as

$$\frac{\partial f_i}{\partial t} + (\mathbf{u} \cdot \nabla) f_i + (\mathbf{F} \cdot \nabla_p) f_i = \left( \frac{Df_i}{Dt} \right)_{\text{coll}} , \quad (1)$$

where  $\nabla$  and  $\nabla_p$  are the usual nabla differential operators with respect to position and momentum components, respectively;  $\mathbf{u}$  is the particle velocity, and  $\mathbf{F}$  is the external force. The term  $(Df/Dt)_{\text{coll}}$  is the so-called collisional term, which describes creations and destructions of particles of type  $i$  with the position  $(\mathbf{r}, \mathbf{r} + d\mathbf{r})$  and momentum  $(\mathbf{p}, \mathbf{p} + d\mathbf{p})$ .

The kinetic equation provides a full description of the system. However, the number of unknowns is enormous. It should be realized that the individual particles are, in general, not just the atoms and ions, but in fact all the individual excitation states of atoms, ions, or molecules. According to the standard procedure, one simplifies the system by constructing equations for the *moments* of the distribution function, i.e. the integrals over momentum weighted by various powers of  $\mathbf{p}$ . I shall present only the final equations; the reader is referred to any standard textbook of the kinetic theory for a detailed derivation and an extensive discussion.

The resulting set of equations are the well-known hydrodynamic equations, namely the continuity equation,

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0 , \quad (2)$$

the momentum equation,

$$\frac{\partial(\rho\mathbf{v})}{\partial t} + \nabla \cdot (\rho\mathbf{v}\mathbf{v}) = -\nabla P + \mathbf{f} \quad , \quad (3)$$

and the energy balance equation,

$$\frac{\partial}{\partial t} \left( \frac{1}{2}\rho v^2 + \rho\epsilon \right) + \nabla \cdot \left[ \left( \frac{1}{2}\rho v^2 + \rho\epsilon + P \right) \mathbf{v} \right] = \mathbf{f} \cdot \mathbf{v} - \nabla \cdot (\mathbf{F}_{\text{rad}} + \mathbf{F}_{\text{con}}) \quad . \quad (4)$$

Here,  $\mathbf{v}$  is the macroscopic velocity,  $\rho$  the total mass density,  $P$  the pressure,  $\mathbf{f}$  the external force,  $\epsilon$  the internal energy,  $\mathbf{F}_{\text{rad}}$  the radiation flux, and  $\mathbf{F}_{\text{con}}$  the conductive flux. Equations (2) - (4) represent moment equations of the kinetic equation, (1), summed over all kinds of particles.

We may also write a zeroth-order moment equation for the individual kinds of particles, i.e. the conservation equation for particles of type  $i$ ,

$$\frac{\partial n_i}{\partial t} + \nabla \cdot (n_i \mathbf{v}) = \left( \frac{Dn_i}{Dt} \right)_{\text{coll}} \quad , \quad (5)$$

where  $n_i$  is the number density (or the occupation number, or population) of particles of type  $i$ . One may also write momentum and energy balance equations for the individual particles if needed (e.g. if different kinds of particles have different macroscopic velocities). We will not consider these situations here.

The moment equations are still quite general. An application of those equations is discussed in other papers of this volume. Here, I will present a further significant simplification of the system, which applies for the case of a stationary (i.e.  $\partial/\partial t = 0$ ), and moreover static ( $\mathbf{v} = 0$ ) medium. Finally, we will consider a 1-D situation, i.e. all quantities depend on the  $z$ -coordinate only,

$$\left( \frac{Dn_i}{Dt} \right)_{\text{coll}} = 0 \quad , \quad (6)$$

$$\nabla P = \mathbf{f} \quad \implies \quad \frac{dP}{dz} = -\rho g \quad , \quad (7)$$

$$\nabla F_{\text{rad}} = 0 \quad \implies \quad F_{\text{rad}} = \text{const} \equiv \sigma T_{\text{eff}}^4 \quad , \quad (8)$$

where  $\sigma$  is the Stefan-Boltzmann constant, and  $T_{\text{eff}}$  is the so-called effective temperature. The first equation is called the *statistical equilibrium* equation, the second one the *hydrostatic equilibrium* equation, while the last one, expressing the fact that the only mechanism that transports energy is radiation, is called the *radiative equilibrium* equation. (Notice that the conductive flux was neglected here, which is a common approximation in the stellar atmospheres theory. However, this approximation breaks down, for instance, in the solar transition region.)

What about convection, which we know may contribute significantly to the energy balance in certain types of stellar atmospheres? Convection is a

transport of energy by rising and falling bubbles of material with properties (e.g. temperature) different from the ambient medium. It is therefore, by its very nature, a non-stationary and non-homogeneous phenomenon. Putting  $\mathbf{v} = 0$  and assuming a 1-D medium means, strictly speaking, that convection is a priori neglected. However, there are descriptions, like the mixing-length theory (see any standard textbook, like Mihalas 1978), that simplify the problem and cast it in the form of 1-D stationary equation, viz.

$$F_{\text{rad}} + F_{\text{conv}} = \sigma T_{\text{eff}}^4 \quad , \quad (9)$$

where the convective flux  $F_{\text{conv}}$  is a specified function of basic state parameters (temperature, density, etc.)

So far, we have specified the kinetic equation for particles. The same may be done for photons. Since, as explained above, photons have a special significance in stellar atmospheres, we will consider the kinetic equation for photons – the so-called *radiative transfer equation* – in the Sect. 2.

### 1.3 LTE versus non-LTE

It is well known from statistical physics that a description of material properties is greatly simplified if the thermodynamic equilibrium (TE) holds. In this state, the particle velocity distributions as well as the distributions of atoms over excitation and ionization states are specified uniquely by two thermodynamic variables. In the stellar atmospheres context, these variables are usually chosen to be the absolute temperature  $T$ , and the total particle number density,  $N$ , or the electron number density,  $n_e$ . From the very nature of a stellar atmosphere it is clear that it cannot be in thermodynamic equilibrium – we *see* a star, therefore we know that photons must be escaping. Since photons carry a significant momentum and energy, the elementary fact of photon escape has to give rise of significant *gradients* of the state parameters in the stellar outer layers.

However, even if the assumption of TE cannot be applied for a stellar atmosphere, we may still use the concept of *local thermodynamic equilibrium* – LTE. This assumption asserts that we may employ the standard thermodynamic relations not globally for the whole atmosphere, but *locally*, for local values of  $T(\mathbf{r})$  and  $N(\mathbf{r})$  or  $n_e(\mathbf{r})$ , despite the gradients that exist in the atmosphere. This assumption simplifies the problem enormously, for it implies that all the particle distribution functions may be evaluated locally, without reference to the physical ensemble in which the given material is found. Notice that the equilibrium values of distribution functions are assigned to *massive particles*; the radiation field is allowed to depart from its equilibrium, Planckian – (22), distribution function.

Specifically, LTE is characterized by the following three distributions:  
– Maxwellian velocity distribution of particles

$$f(\mathbf{v})d\mathbf{v} = (m/2\pi kT)^{3/2} \exp(-mv^2/2kT) d\mathbf{v} \quad , \quad (10)$$

where  $m$  is the particle mass, and  $k$  the Boltzmann constant.

– Boltzmann excitation equation,

$$(n_j/n_i) = (g_j/g_i) \exp[-(E_j - E_i)/kT] \quad , \quad (11)$$

where  $g_i$  is the statistical weight of level  $i$ , and  $E_i$  the level energy, measured from the ground state.

– Saha ionization equation,

$$\frac{N_I}{N_{I+1}} = n_e \frac{U_I}{U_{I+1}} C T^{-3/2} \exp(\chi_I/kT) \quad , \quad (12)$$

where  $N_I$  is the total number density of ionization stage  $I$ ,  $U$  is the partition function, defined by  $U = \sum_1^\infty g_i \exp(-E_i/kT)$ ;  $\chi_I$  is the ionization potential of ion  $I$ , and  $C = (h^2/2\pi mk)^{3/2}$  is a constant ( $= 2.07 \times 10^{-16}$  in cgs units). It should be stressed that in the astrophysical LTE description, the *same* temperature  $T$  applies to all kinds of particles, and to all kinds of distributions, (10) – (12).

Equations (10) – (12) define the state of LTE from the macroscopic point of view. Microscopically, LTE holds if all atomic processes are in *detailed balance*. i.e. if the number of processes  $A \rightarrow B$  is exactly balanced by the number of inverse processes  $B \rightarrow A$ . By  $A$  and  $B$  we mean any particle states between which there exists a physically reasonable transition. For instance,  $A$  is an atom in an excited state  $i$ , and  $B$  the same atom in another state  $j$  (either of the same ion as  $i$ , in which case the process is an excitation/de-excitation; or of the higher or lower ion, in which case the term is an ionization/recombination). An illuminating discussion is presented in the textbook by Oxenius (1986).

In contrast, by the term non-LTE (or NLTE) we understand any state that departs from LTE. In practice, one usually means that populations of some selected energy levels of some selected atoms/ions are allowed to depart from their LTE value, while velocity distributions of all particles are assumed to be Maxwellian, (10), moreover with the same kinetic temperature,  $T$ .

One of the big issues of modern stellar atmospheres theory is whether, and if so to what extent, should departures from LTE be accounted for in numerical modeling. This question will be discussed in more detail later on (Sects. 3, 5). Generally, to understand why and where we may expect departures from LTE, let us turn to the microscopic definition of LTE. It is clear that LTE breaks down if the detailed balance in at least one transition  $A \rightarrow B$  breaks down. We distinguish the *collisional transitions* (arising due to interactions between two or more massive particles), and *radiative transitions* (interactions involving particles and photons). Under stellar atmospheric conditions, collisions between massive particles tend to maintain the local equilibrium (since velocities are Maxwellian). Therefore, the validity of LTE hinges on whether the radiative transitions are in detailed balance or not.

Again, the fact that the radiation escapes from a star implies that LTE should eventually break down at a certain point in the atmosphere. Essentially, this is because detailed balance in radiative transitions generally breaks down at a certain point near the surface. Since photons escape (and more so from the uppermost layers), there must be a lack of them there. Consequently, the number of photoexcitations (or any atomic transition induced by absorbing a photon) is less than a number of inverse processes, spontaneous de-excitations (we neglect here, for simplicity, stimulated emission).

These considerations explain that we may expect departures from LTE if the following two conditions are met: i) radiative rates in some important atomic transition dominate over the collisional rates; and ii) radiation is not in equilibrium, i.e. the intensity does not have the Planckian distribution. Later, we will show how these conditions are satisfied in different stellar types. However, some general features can be seen immediately. Collisional rates are proportional to the particle density; it is therefore clear that for high densities the departures from LTE tend to be small. Likewise, deep in the atmosphere, photons do not escape, and so the intensity is close to the equilibrium value. Departures from LTE are therefore small, even if the radiative rates dominate over the collisional rates.

## 2 Radiative Transfer Equation

As explained above, radiation plays a somewhat privileged role in the stellar atmospheres theory. This is the reason why we consider the radiative transfer equation separately from equations describing the material properties. The dominant role of radiation is also reflected in the terminology – the whole stellar atmosphere problem is sometimes referred to as a *solving the radiative transfer equation with constraints*, i.e. viewing all the material equations as mere “constraints”.

As discussed in the preceding section, one may view the radiative transfer equation as a kinetic equation for photons. In the astronomical literature, it is customary to start with a phenomenological derivation of the radiative transfer equation, and to show later that this equation is in fact equivalent to the kinetic equation.

It should be realized, however, that when viewing radiation as an ensemble of mutually non-interacting, massless particles – photons, and describing the interaction between radiation and matter in terms of simple collisions (interactions) between photons and massive particles, the wave phenomena connected with radiation are in fact neglected. This is a good approximation if i) the wavelength of radiation is much smaller than the typical distance between massive particles; and if ii) the particle positions are random. These conditions are well satisfied under the stellar atmospheric conditions: we deal with a hot plasma, so the particle positions are indeed random. For optical, UV, and even higher-frequency radiation, the wavelengths ( $\lambda < 10^{-4}$  cm)

are indeed smaller than typical interparticle distances. For infrared and radio wavelengths, some wave phenomena (e.g. refraction) may actually play a role in the radiative transfer.

In the following, we will adhere to the photon picture, and neglect all the wave phenomena. A somewhat special case is the polarization of radiation. Polarization will also be neglected here, i.e. we assume an unpolarized radiation. We will see in other lectures (e.g., Bjorkman, this volume), that polarization of radiation may actually play an important diagnostic role in certain stellar atmospheric structures (as an indicator of asymmetries of the medium). It is possible to extend the usual formalism of the transfer equation to account for polarization, by introducing a vector quantity (the so-called Stokes vector) instead of scalar intensity of radiation, and to write down the transfer equation in the vector form. We will not consider this case here; the interested reader is referred to standard textbooks – Chandrasekhar (1960); or recently Stenflo (1994); or excellent review articles (several papers in Kalkofen 1987).

## 2.1 Intensity of Radiation and Related Quantities

We start with phenomenological definitions. The *specific intensity*,  $I(\mathbf{r}, \mathbf{n}, \nu, t)$ , of radiation at position  $\mathbf{r}$ , traveling in direction  $\mathbf{n}$ , with frequency  $\nu$ , at time  $t$  is defined such that the energy transported by radiation in the frequency range  $(\nu, \nu + d\nu)$ , across an elementary area  $dS$ , into a solid angle  $d\omega$  in a time interval  $dt$  is

$$dE = I(\mathbf{r}, \mathbf{n}, \nu, t) dS \cos \theta d\omega d\nu dt , \quad (13)$$

where  $\theta$  is the angle between  $\mathbf{n}$  and the normal to the surface  $dS$  (i.e.  $dS \cos \theta = \mathbf{n} \cdot d\mathbf{S}$ ). The dimension of  $I$  is  $\text{erg cm}^{-2} \text{sec}^{-1} \text{hz}^{-1} \text{sr}^{-1}$ . The specific intensity provides a complete description of the unpolarized radiation field from the macroscopic point of view.

As pointed out above, there is a close connection between the specific intensity and the photon distribution function,  $f$ . The latter is defined such that  $f(\mathbf{r}, \mathbf{n}, \nu, t) d\omega d\nu$  is the number of photons per unit volume at location  $\mathbf{r}$  and time  $t$ , with frequencies in the range  $(\nu, \nu + d\nu)$ , propagating with velocity  $c$  in direction  $\mathbf{n}$ . The number of photons crossing an element  $d\mathbf{S}$  in time  $dt$  is  $f(c \cdot dt)(\mathbf{n} \cdot d\mathbf{S})(d\omega d\nu)$ . The energy of those photons is the same expression multiplied by  $h\nu$ ,  $h$  being the Planck constant. Comparing this to the definition of the specific intensity, we obtain the desired relation between the specific intensity and the distribution function,

$$I = (c h \nu) f . \quad (14)$$

This relation makes it easy to understand the following expressions. Analogously as for massive particles, one defines various moments of the distribution function – i.e. the specific intensity in this case – which have a meaning of the energy density, flux, and the stress tensor.

From the definition of the distribution function, it is clear that the *energy density* of radiation is given by (dropping an explicit indication of the dependence on frequency, etc.)

$$E = \oint (h\nu) f \, d\omega = (1/c) \oint I \, d\omega \quad , \quad (15)$$

because  $f$  is the number of photons in an elementary volume, and  $h\nu$  the energy of each; we have to integrate over all solid angles. Similarly, the *energy flux* of radiation is given by

$$\mathbf{F} = \oint (h\nu) \cdot (c \mathbf{n}) f \, d\omega = \oint \mathbf{n} I \, d\omega \quad , \quad (16)$$

because  $c \mathbf{n}$  is the vector velocity.

The *radiation stress tensor* is defined by

$$\mathbf{P} = \oint (h\nu) \mathbf{n} \mathbf{n} f \, d\omega = (1/c) \oint \mathbf{n} \mathbf{n} I \, d\omega \quad . \quad (17)$$

Finally, we mention that the *photon momentum density* [recall that the momentum of an individual photon is  $(h\nu/c)\mathbf{n}$ ] is given by

$$\mathbf{G} = \oint (h\nu/c) \mathbf{n} f \, d\omega = (1/c^2) \mathbf{F} \quad , \quad (18)$$

i.e., it is proportional to the radiation flux.

## 2.2 Absorption and Emission Coefficient

The radiative transfer equation describes the changes of the radiation field due to its interaction with matter. To describe this interaction, one first introduces several phenomenological quantities:

*Absorption coefficient* describes the removal of energy from the radiation field by matter. It is defined in such a way that an element of material, of cross-section  $dS$  and length  $ds$ , removes from a beam of specific intensity  $I$  (incident normal to  $dS$  into a solid angle  $d\omega$ ), an amount of energy

$$dE = \chi(\mathbf{r}, \mathbf{n}, \nu, t) I(\mathbf{r}, \mathbf{n}, \nu, t) \, dS \, d\omega \, d\nu \, dt \quad . \quad (19)$$

The dimension of  $\chi$  is  $\text{cm}^{-1}$ , thus  $1/\chi$  has a dimension of length, and it measures a characteristic distance a photon can travel before it is absorbed; in other words, the *photon mean free path*.

*Emission coefficient* describes the energy released by the material in the form of radiation. Analogously, it is defined such as an elementary volume of material, of cross-section  $dS$  and length  $ds$ , releases (into a solid angle  $d\omega$ , in direction  $\mathbf{n}$ , within a frequency band  $d\nu$ ) an amount of energy

$$dE = \eta(\mathbf{r}, \mathbf{n}, \nu, t) \, dS \, d\omega \, d\nu \, dt \quad . \quad (20)$$

The dimension of  $\eta$  is  $\text{erg cm}^{-3} \text{hz}^{-1} \text{sec}^{-1} \text{sr}^{-1}$ .

The absorption and emission coefficients are defined *per unit length*. Sometimes, one defines the coefficients *per unit mass*, which are given by expressions (19) and (20) divided by the mass density,  $\rho$ .

The above coefficients are defined phenomenologically. To be able to write down actual expressions for them, we have to go to microscopic physics. In other words, we have to describe all contributions from microscopic processes that give rise to an absorption or emission of photons with specified properties. Detailed expressions will be considered later on (Sects. 3.1, 5.2). Here, we will discuss some general points:

i) Sometimes, one distinguishes two types of absorption, a “true absorption”, and a “scattering”. In the true absorption (also called “thermal absorption”) process, a photon is removed from the incident beam and is destroyed; while in the scattering process a photon is first removed from the beam, but is immediately re-emitted in a different direction and with (slightly) different frequency. This distinction is reflected in a notation,

$$\chi(\mathbf{r}, \mathbf{n}, \nu, t) = \kappa(\mathbf{r}, \mathbf{n}, \nu, t) + \sigma(\mathbf{r}, \mathbf{n}, \nu, t) , \quad (21)$$

where the first term on the right hand side,  $\kappa$ , refers to the true absorption, while the second term,  $\sigma$ , to the scattering. However, I stress that this distinction does not really have to do much with the absorption process –  $\chi$  describes a removal of photon from the beam and does not have to care about what happens next. The distinction between the true absorption and scattering actually enters rather the proper formulation of the emission coefficient.

ii) It is known from the quantum theory of radiation that there are three types of elementary processes that give rise to an absorption or emission of a photon: 1) induced absorption – an absorption of a photon accompanied by a transition of an atom/ion to a higher energy state); 2) spontaneous emission – an emission of a photon accompanied by a spontaneous transition of an atom/ion to a lower energy state); and 3) stimulated emission – an interaction of an atom/ion with a photon accompanied by an emission of another photon with identical properties. In the astrophysical formalism, the stimulated emission is usually treated as negative absorption.

iii) In thermodynamic equilibrium, the microscopic detailed balance holds, and therefore the radiation energy absorbed in an elementary volume in an elementary frequency interval is exactly balanced by the energy emitted in the same volume and in the same frequency range. From the definition expressions for the absorption and emission coefficients, (19) and (20), it follows that in the equilibrium state,  $\chi I = \eta$ . Moreover, we know that in thermodynamic equilibrium the radiation intensity is equal to the Planck function,  $I = B$ , where

$$B(\nu, T) = \frac{2h\nu^3}{c^2} \frac{1}{\exp(h\nu/kT) - 1} . \quad (22)$$

We are then left with an interesting relation that in thermodynamic equilibrium,  $\eta/\chi = B$ , which is called *Kirchhoff's law*.

### 2.3 Phenomenological Derivation of the Transfer Equation

Having defined the basic phenomenological coefficients which describe the interaction of radiation and matter, a heuristic derivation of the radiative transfer equation is straightforward. We express a conservation of the total photon energy when a radiation beam passes through an elementary volume of matter of cross-section  $dS$  (perpendicular to the direction of propagation) and length  $ds$  (measured along the direction of propagation). Taking into account definitions of the specific intensity, (13), and the absorption and emission coefficients, (19) and (20), we obtain

$$\begin{aligned} & [I(\mathbf{r} + \Delta\mathbf{r}, \mathbf{n}, \nu, t + \Delta t) - I(\mathbf{r}, \mathbf{n}, \nu, t)] dS d\omega d\nu dt = \\ & [\eta(\mathbf{r}, \mathbf{n}, \nu, t) - \chi(\mathbf{r}, \mathbf{n}, \nu, t)I(\mathbf{r}, \mathbf{n}, \nu, t)] ds dS d\omega d\nu dt \quad , \end{aligned} \quad (23)$$

which expresses the fact that the difference between specific intensities before and after passing through the elementary volume of pathlength  $ds$  is equal to the difference of the energy emitted and absorbed in the volume. The difference of intensities on the left hand side may be expressed as

$$\text{diff } I = \frac{\partial I}{\partial s} ds + \frac{\partial I}{\partial t} dt = \left( \frac{\partial I}{\partial s} + \frac{1}{c} \frac{\partial I}{\partial t} \right) ds \quad . \quad (24)$$

Finally,  $\partial I/\partial s$  may be written as  $\mathbf{n} \cdot \nabla$ , so we arrive at

$$\left( \frac{1}{c} \frac{\partial}{\partial t} + \mathbf{n} \cdot \nabla \right) I(\mathbf{r}, \mathbf{n}, \nu, t) = \eta(\mathbf{r}, \mathbf{n}, \nu, t) - \chi(\mathbf{r}, \mathbf{n}, \nu, t) I(\mathbf{r}, \mathbf{n}, \nu, t) \quad . \quad (25)$$

This is the general form of the radiative transfer equation. Let us now consider two important special cases.

1) for a one-dimensional planar atmosphere,  $n_z = (dz/ds) = \cos \theta \equiv \mu$ , where  $\theta$  is the angle between direction of propagation of radiation,  $\mathbf{n}$ , and the normal to the surface. Further, let us assume a time-independent situation,  $\partial/\partial t = 0$ , so we obtain

$$\mu \frac{dI(\nu, \mu, z)}{dz} = \eta(\nu, \mu, z) - I(\nu, \mu, z) \chi(\nu, \mu, z) \quad , \quad (26)$$

where the intensity of radiation is now only a function of the geometrical coordinate  $z$ , frequency  $\nu$ , and the directional cosine  $\mu$ .

2) in spherical coordinates, the derivative along the ray,  $\partial/\partial s$  is given by  $\partial/\partial s = \mu(\partial/\partial r) + (1 - \mu^2)/r(\partial/\partial \mu)$ , and the radiative transfer equation in a spherically symmetric medium is written as

$$\mu \frac{\partial I(\nu, \mu, r)}{\partial r} + \frac{1 - \mu^2}{r} \frac{\partial I(\nu, \mu, r)}{\partial \mu} = \eta(\nu, \mu, r) - I(\nu, \mu, r) \chi(\nu, \mu, r) \quad . \quad (27)$$

## 2.4 Optical Depth and the Source Function

Let us start with a simple 1-D transfer equation, written as

$$\mu \frac{dI_\nu}{dz} = \eta_\nu - \chi_\nu I_\nu \quad , \quad (28)$$

where we drop an explicit indication of the dependence of  $I$ ,  $\eta$ , and  $\chi$  on the geometrical distance  $z$  and angle  $\mu$ , and write, as is customary in the astrophysical literature, the frequency  $\nu$  as a subscript. We divide the transfer equation by  $\chi_\nu$ , and obtain a very simple and advantageous form of the transfer equation,

$$\mu \frac{dI_\nu}{d\tau_\nu} = I_\nu - S_\nu \quad , \quad (29)$$

where the *elementary optical depth* is defined by

$$d\tau_\nu \equiv -\chi_\nu dz \quad , \quad (30)$$

and the *source function* is defined by

$$S_\nu \equiv \frac{\eta_\nu}{\chi_\nu} \quad . \quad (31)$$

The absorption and emission coefficient are *local* quantities, therefore the definition of the source function, (31), applies for all geometries. The optical depth depends on the geometry; in case of a 3-D transfer, the most natural definition is the *optical depth along the ray*, defined by

$$d\tau = \chi(\mathbf{r}, \mathbf{n}, \nu) ds \quad , \quad (32)$$

where  $ds$  is the elementary pathlength in the direction  $\mathbf{n}$ . In the case of a plane-parallel atmosphere, the relation between the optical depth in the direction  $\mu$  (which we denote here as  $\tau_{\mu\nu}$ ), and the “normal-direction”  $\tau_\nu$  defined by (30), is

$$d\tau_{\mu\nu} = d\tau_\nu / \mu \quad . \quad (33)$$

What is the physical meaning of the optical depth and of the source function? The meaning of the optical depth is straightforward. In the absence of emissions, the transfer equation is simply  $dI/d\tau = I$ , and the solution is  $I(\tau) = I(\tau + \Delta\tau) \exp(-\Delta\tau)$ , i.e. the optical depth is the e-folding distance for attenuation of the specific intensity due to absorption. In other words, the probability that a photon will travel an optical distance  $\tau$  is simply  $p(\tau) = \exp(-\tau)$ . Since the absorption coefficient (e.g. in spectral lines) may be a sharply varying function of frequency, the (monochromatic) optical depth may also vary significantly with frequency. Sometimes one defines various frequency-independent optical depths, like those corresponding to the averaged absorption coefficient, either over the whole spectrum (a typical example being the Rosseland optical depth – see Sect. 2.8), or over a spectral line (see Sect. 3.1).

The meaning of the source function can also be easily understood. Let us write the number of photons emitted in an elementary volume (defined by an elementary area  $dS$  and an elementary path  $ds$ ), to all directions. From the definition of the emission coefficient it follows that (assuming an isotropic emission for simplicity)  $N_{\text{em}} = \eta ds (4\pi/h\nu) d\nu dt dS$ , where the factor  $4\pi$  comes from an integration over all solid angles, and  $h\nu$  transforms energy (from the original definition of the emission coefficient) to the number of photons. Using the definition of the optical depth and the source function, we may rewrite the factor  $\eta ds$  as  $\eta ds = (\eta/\chi)\chi ds = S(\tau)d\tau$ . Consequently, the number of emitted photons is

$$N_{\text{em}} = S(\tau)d\tau \frac{4\pi}{h\nu} d\nu dt dS . \quad (34)$$

In other words, the source function is proportional to the *number of photons emitted per unit optical depth interval*.

For completeness, we mention that the number of photons *absorbed* per unit optical depth interval (from all solid angles) is analogously given by

$$N_{\text{abs}} = J(\tau)d\tau \frac{4\pi}{h\nu} d\nu dt dS , \quad (35)$$

which directly follows from (19);  $J$  being the mean intensity of radiation, defined by (48).

## 2.5 Elementary Solutions

In this section, we consider the simplest solutions of the 1-D plane-parallel transfer equation. For notational simplicity, we drop subscript  $\nu$  indicating the frequency dependence.

a) *No absorption, no emission*, i.e.  $\chi = \eta = 0$ . The transfer equation reads  $dI/dz = 0$ , which has a trivial solution

$$I = \text{const} . \quad (36)$$

This expresses the obvious fact that in the absence of any interaction with the medium, the radiation intensity remains constant.

b) *No absorption, only emission*, i.e.  $\chi = 0$ , but  $\eta > 0$ . The solution is simply

$$I(z, \mu) = I(0, \mu) + \int_0^z \eta(z') dz' / \mu . \quad (37)$$

This equation is often used for describing an outgoing radiation from an optically thin radiating slab, like for instance a forbidden line radiation from planetary nebulae, or a radiation from the solar transition region and/or corona.

c) *No emission, only absorption*, i.e.  $\eta = 0$ ,  $\chi > 0$ . The transfer equation now reads  $\mu dI/d\tau = I$ , and the solution is simply

$$I(0, \mu) = I(\tau, \mu) \exp(-\tau/\mu) . \quad (38)$$

d) *Absorption and emission*. We will now write a general formal solution of the transfer equation, i.e. for the case where both, absorption and emission, coefficients are different from zero,  $\chi > 0$ ,  $\eta > 0$ . The solution is called “formal” because it is assumed here that both  $\chi$  and  $\eta$  are *specified* functions of position and frequency. As we shall see later on, both coefficients may depend on the radiation field, so that in actual problems they may not be given a priori, without previously solving the general transfer problem. The formal solution reads

$$I(\tau_1, \mu) = I(\tau_2, \mu) \exp[-(\tau_2 - \tau_1)/\mu] + \int_{\tau_1}^{\tau_2} S(t) \exp[-(t - \tau_1)/\mu] dt/\mu . \quad (39)$$

e) *Semi-infinite atmosphere*. A special case of the formal solution (39) for emergent radiation (i.e.  $\tau_1 = 0$ ) from a semi-infinite atmosphere ( $\tau_2 = \infty$ ) reads

$$I(0, \mu) = \int_0^{\infty} S(t) \exp(-t/\mu) dt/\mu . \quad (40)$$

This equation shows that the specific intensity in a semi-infinite atmosphere is in fact a Laplace transform of the source function.

f) *Semi-infinite atmosphere with a linear source function*. Another special case of the general formal solution (39) is a emergent intensity from a semi-infinite atmosphere, with a source function being a *linear* function of optical depth,  $S(\tau) = a + b\tau$ . It is given by

$$I(0, \mu) = a + b\mu = S(\tau = \mu) . \quad (41)$$

This important expression is called the *Eddington-Barbier relation*. It shows that the emergent intensity, for instance in the normal direction ( $\mu = 1$ ) is given by the value of the source function at the optical depth of unity. The values of emergent intensity for all angles  $\mu$  between 0 and 1 then map the values of the source function between optical depths 0 and 1. Although in reality the source function does not have to be a linear function of optical depth, it can usually be well approximated by it in the vicinity of  $\tau = 1$ . Consequently, the Eddington-Barbier relation, (41), usually provides a good estimate of the emergent intensity.

g) *Finite homogeneous slab*. Finally, an expression for an emergent radiation ( $\tau_1 = 0$ ) from a finite ( $\tau_2 = T < \infty$ ) and homogeneous slab [i.e.  $S(t) = S$  is constant], in the normal direction ( $\mu = 1$ ), reads

$$I(0, 1) = S \cdot (1 - e^{-T}) . \quad (42)$$

In the special case  $T \gg 1$ , (42) becomes  $I(0, 1) = S$ , while for  $T \ll 1$ , we obtain  $I(0, 1) = S \cdot T$ . Both limiting expressions have a simple physical explanation. As we have shown above, the source function expresses a number of photons (or radiative energy) emitted per unit optical depth. In the optically thin case ( $T \ll 1$ ), there is little absorption, so practically all created photons escape from the medium. Since the actual optical depth is  $T$ , the total emergent intensity is  $S \cdot T$ . In the optical thick case, we may roughly say that the photons created deeper than  $\tau = 1$  are very likely absorbed, so the only photons which contribute to the emergent intensity are those emitted at optical depths  $\tau \leq 1$ . Consequently, the emergent intensity is  $S \cdot 1$ , i.e.  $S$ , regardless of the actual optical thickness of the slab.

## 2.6 Moments of the Transfer Equation

Analogously to the case of massive particles, we may define various moments of the photon distribution function, i.e. the specific intensity. By appropriately integrating the kinetic (i.e. transfer) equation we obtain relations between these moments. As was discussed in Sect. 2.1., the first three moments are the photon energy density, radiation flux, and the radiation stress tensor. Written synoptically, (see, e.g. the textbook by Shu 1991), we may write

$$\begin{pmatrix} cE_\nu \\ \mathbf{F}_\nu \\ c\mathbf{P}_\nu \end{pmatrix} = \oint \begin{pmatrix} 1 \\ \mathbf{n} \\ \mathbf{nn} \end{pmatrix} I_\nu d\omega . \quad (43)$$

Consequently, the moment equations are obtained by multiplying the transfer equation (25) by 1,  $\mathbf{n}$ , etc., and integrating over all solid angles. The first two moment equations read

$$\frac{\partial E_\nu}{\partial t} + \nabla \cdot \mathbf{F}_\nu = \eta_\nu - \chi_\nu c E_\nu , \quad (44)$$

$$\frac{1}{c} \frac{\partial \mathbf{F}_\nu}{\partial t} + c \nabla \cdot \mathbf{P}_\nu = -\chi_\nu \mathbf{F}_\nu , \quad (45)$$

Both these equations have the general structure of the moment equation of the kinetic equation, namely

$$\partial/\partial t(\text{density of quantity}) + (\text{gradient of its flux}) = (\text{sources} - \text{sinks}) . \quad (46)$$

In the astrophysical literature, one usually introduces moments as *angle-averaged*, rather than angle-integrated quantities. The first moments of the radiation intensity are usually denoted as  $J$ ,  $H$ ,  $K$ . We may synoptically write

$$\begin{pmatrix} J_\nu \\ \mathbf{H}_\nu \\ \mathbf{K}_\nu \end{pmatrix} = \frac{1}{4\pi} \begin{pmatrix} cE_\nu \\ \mathbf{F}_\nu \\ c\mathbf{P}_\nu \end{pmatrix} = \frac{1}{4\pi} \oint \begin{pmatrix} 1 \\ \mathbf{n} \\ \mathbf{nn} \end{pmatrix} I_\nu d\omega . \quad (47)$$

In a plane-parallel approximation, all the moments are scalar quantities, and are given by

$$J_\nu = \frac{1}{2} \int_{-1}^1 I_\nu(\mu) d\mu , \quad (48)$$

$$H_\nu = \frac{1}{2} \int_{-1}^1 \mu I_\nu(\mu) d\mu , \quad (49)$$

$$K_\nu = \frac{1}{2} \int_{-1}^1 \mu^2 I_\nu(\mu) d\mu , \quad (50)$$

and the moment equations are written as

$$\frac{dH_\nu}{d\tau_\nu} = J_\nu - S_\nu , \quad (51)$$

and

$$\frac{dK_\nu}{d\tau_\nu} = H_\nu . \quad (52)$$

The system of moment equations is not closed, i.e. the equation for  $n$ -th moment contains the  $(n+1)$ -th moment, etc. It is therefore necessary to come up with some kind of closure relation. In the stellar atmospheres theory, one defines the so-called Eddington factor,  $f^K$ , by

$$f_\nu^K \equiv K_\nu / J_\nu . \quad (53)$$

It is clear from the definition of moments that in the case of isotropic radiation,  $I_\nu(\mu) = I_\nu$  being independent of angle, the Eddington factor  $f^K = 1/3$ . Assuming the Eddington factor to be specified, one may combine the two moment equations (51) and (52) together,

$$\frac{d^2(f_\nu^K J_\nu)}{d\tau_\nu^2} = J_\nu - S_\nu . \quad (54)$$

This equation is very useful. It effectively eliminates one independent variable, the angle  $\mu$ , from the problem. Numerically, it replaces the original transfer equation, which is a first-order linear differential equation for the specific intensity  $I_{\nu\mu}$ , by a second-order but still linear differential equation for the mean intensity,  $J_\nu$ . However, its simplicity is illusory. It cannot be used alone, even if the source function is given, since in general the Eddington factor is unknown unless the full solution of the transfer equation is known. However, this form of the transfer equation is very useful in certain numerical methods, as we will discuss later on. It should be realized that it can be used to advantage only in *iterative* methods; in which we use current values of  $J$  and  $K$  to determine the current Eddington factor  $f^K$ , and keep this factor fixed during the subsequent iteration step.

## 2.7 Lambda Operator

Let us first write down the general formal solution of the radiative transfer equation for a semi-infinite plane-parallel atmosphere, with no incoming radiation at the surface ( $\tau = 0$ ),

$$I_\nu(\tau_\nu, \mu) = \int_{\tau_\nu}^{\infty} S_\nu(t) e^{-(t-\tau_\nu)/\mu} dt / \mu, \quad \text{for } \mu \geq 0, \quad (55)$$

$$I_\nu(\tau_\nu, \mu) = \int_0^{\tau_\nu} S_\nu(t) e^{-(\tau_\nu-t)/(-\mu)} dt / (-\mu), \quad \text{for } \mu < 0. \quad (56)$$

Recall that from the definition of the directional cosine  $\mu$  follows that positive values of  $\mu$  correspond to outward directions, while negative values of  $\mu$  correspond to inward directions. The mean intensity of radiation is obtained by integrating (55) and (56) over  $\mu$ , viz.

$$J_\nu(\tau_\nu) = \frac{1}{2} \int_0^{\infty} S_\nu(t) E_1(|t - \tau_\nu|) dt, \quad (57)$$

where  $E_1$  is the first exponential integral. The general exponential integral is defined by

$$E_n(x) \equiv \int_1^{\infty} \frac{e^{-xt}}{t^n} dt, \quad (58)$$

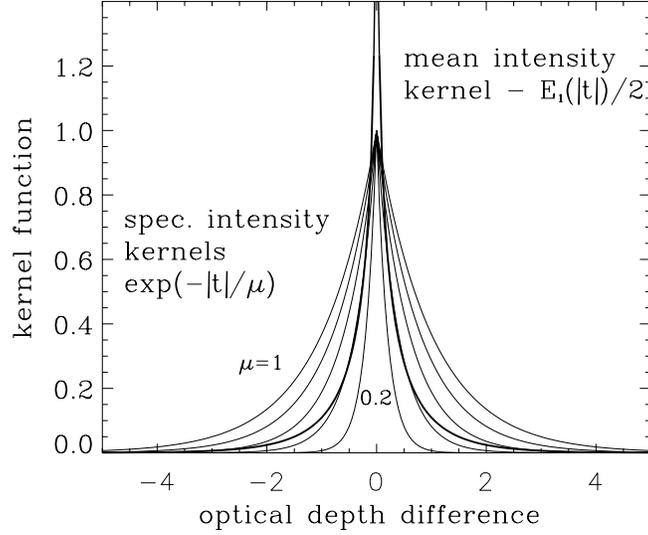
The mean intensity may be synoptically expressed as an action of an operator,  $A$ , on the source function,

$$J_\nu(\tau_\nu) = A_{\tau_\nu} [S(t)], \quad (59)$$

where the  $A$ -operator is defined by

$$A_\tau [f(t)] = \frac{1}{2} \int_0^{\infty} E_1(|t - \tau|) f(t) dt, \quad (60)$$

The behavior of the kernel functions corresponding to the specific intensity, (55), which is a simple exponential, and for the mean intensity, which is the first exponential integral, (57), is displayed in Fig. 1 The width of the kernel decreases with decreasing  $\mu$ , which is easily understood by realizing that a unit optical distance for a photon traveling with a certain angle with respect to the normal to the surface,  $\mu$ , corresponds to a larger optical distance than for one traveling in the normal direction, because the distance is proportional to  $1/\mu$ . Similarly, the kernel for the specific intensity propagating in the normal direction ( $\mu = 1$ ) is significantly wider than the kernel for the mean intensity. This is simply because the mean intensity contains contribution from all angles, i.e. the corresponding kernel is an average over all  $\mu$ -dependent specific intensity kernels.



**Fig. 1.** Kernel functions for the specific intensity at various angles (thin lines for  $\mu = 1, 0.8, 0.6, 0.4$ , and  $0.2$ ), and for the mean intensity of radiation (thick line)

For practical purposes, (57) or (59) have to be replaced by a quadrature sum. Equation (59) can thus be written in the discretized form as

$$J_d = \sum_{d'=1}^D \Lambda_{dd'} S_{d'} \quad , \quad (61)$$

where  $d$  denotes the depth index (we dropped the frequency index  $\nu$ ). The  $\Lambda$ -operator can thus be thought of as  $\Lambda$ -matrix, and the mean intensity as well as the source function at all depths as column vectors.

What is the meaning of the  $\Lambda$ -matrix? Let us take, quite formally, all elements of the source function vector to be zero except the  $i$ -th element which is taken to be 1,  $S_d = \delta_{di}$ . Then

$$\begin{pmatrix} J_1 \\ J_2 \\ \vdots \\ J_D \end{pmatrix} = \begin{pmatrix} \Lambda_{11} & \Lambda_{12} & \dots & \Lambda_{1D} \\ \Lambda_{21} & \Lambda_{22} & \dots & \Lambda_{2D} \\ \vdots & \vdots & \ddots & \vdots \\ \Lambda_{D1} & \Lambda_{D2} & \dots & \Lambda_{DD} \end{pmatrix} \times \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \end{pmatrix} = \begin{pmatrix} \Lambda_{1i} \\ \Lambda_{2i} \\ \vdots \\ \Lambda_{Di} \end{pmatrix} \quad . \quad (62)$$

In other words, the  $i$ -th column of the  $\Lambda$  matrix is a solution of the transfer equation with the source function given as a unit pulse function. Physically,

the  $i$ -th column of  $A$  therefore describes how the pulse which originated at the  $i$ -th depth point spreads over all depths.

## 2.8 Diffusion Approximation

Deep in the atmosphere, the source function approaches the Planck function,  $S_\nu \rightarrow B_\nu$ , because virtually no photons escape, and thus the medium approaches the thermal equilibrium. Let us choose a reference optical depth,  $\tau_\nu \gg 1$ , and let us expand the source function for  $t_\nu \geq \tau_\nu$  by a Taylor expansion,

$$S_\nu(t_\nu) = \sum_{n=0}^{\infty} \frac{d^n B_\nu}{d\tau_\nu^n} \frac{(t_\nu - \tau_\nu)^n}{n!} , \quad (63)$$

Substituting this expression to the formal solution, (55) and (56), we obtain

$$I_\nu(t_\nu, \mu) = \sum_{n=0}^{\infty} \mu^n \frac{d^n B_\nu}{d\tau_\nu^n} = B_\nu(\tau_\nu) + \mu \frac{dB_\nu}{d\tau_\nu} + \mu^2 \frac{d^2 B_\nu}{d\tau_\nu^2} + \dots . \quad (64)$$

By substituting this expression into definition equations for the moments, we obtain

$$J_\nu(\tau_\nu) = B_\nu(\tau_\nu) + \frac{1}{3} \frac{d^2 B_\nu}{d\tau_\nu^2} + \dots , \quad (65)$$

$$H_\nu(\tau_\nu) = \frac{1}{3} \frac{dB_\nu}{d\tau_\nu} + \dots , \quad (66)$$

$$K_\nu(\tau_\nu) = \frac{1}{3} B_\nu(\tau_\nu) + \frac{1}{5} \frac{d^2 B_\nu}{d\tau_\nu^2} + \dots . \quad (67)$$

These equations illustrate several features of the behavior of the radiation field at large depths. First, the mean intensity approaches the Planck function. Second, the radiation field is nearly isotropic, and the Eddington factor  $f_\nu^K = K_\nu/J_\nu$  approaches  $1/3$ . Finally, the monochromatic flux is given as a derivative of the Planck function with respect to the optical depth. Since the Planck function is only a function of temperature, we may express the flux by means of the temperature gradient,

$$H_\nu = \frac{1}{3} \frac{dB_\nu}{d\tau_\nu} = -\frac{1}{3} \frac{1}{\chi_\nu} \frac{dB_\nu}{dz} = -\frac{1}{3} \frac{1}{\chi_\nu} \frac{dB_\nu}{dT} \frac{dT}{dz} . \quad (68)$$

Thus, at great depths the transfer problem reduces to this single equation. The name *diffusion approximation* comes from the similarity of this equation to other, material, diffusion equations, which are typically of the form

$$\text{flux} = (\text{diffusion coefficient}) \times (\text{gradient of the relevant quantity}). \quad (69)$$

We may thus think of the term  $(-1/3)(1/\chi_\nu)(dB_\nu/dT)$  as a *radiative diffusion coefficient*; or, because of a similarity of (68) to the heat conductivity equation, as *radiative conductivity*.

By integrating over all frequencies we obtain for the *total radiation flux* in the diffusion approximation

$$H = - \left( \frac{1}{3} \frac{1}{\chi_R} \frac{dB}{dT} \right) \frac{dT}{dz} . \quad (70)$$

where the averaged opacity is defined by

$$\frac{1}{\chi_R} \frac{dB}{dT} = \int_0^\infty \frac{1}{\chi_\nu} \frac{dB_\nu}{dT} d\nu , \quad (71)$$

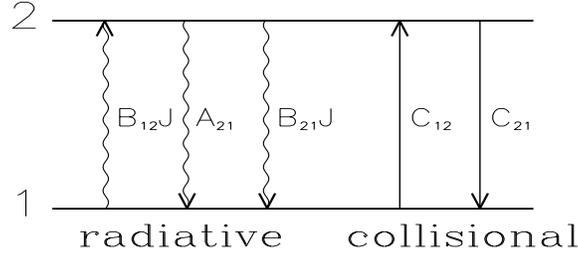
which is the well-known *Rosseland mean opacity*. One may define many other averaged (mean) opacities by simpler expressions, but we see why the Rosseland opacity is defined by this seemingly strange expression – it yields the *exact total radiation flux* at large depths. Since the temperature in the atmosphere is in fact determined by the condition imposed on the total radiation flux, the Rosseland mean opacity yields the *correct temperature structure* deep in the atmosphere. It is also clear why the Rosseland opacity is the most appropriate one for the use in the stellar interior theory (de Greve, this volume). Notice also that the integrand in the definition of Rosseland opacity contains  $1/\chi$ , i.e. the contribution to the integral is largest for the lowest monochromatic opacities. Indeed, for those frequencies the medium is most transparent, and therefore the monochromatic flux is largest. This again shows that the Rosseland mean opacity is the most appropriate one for describing the total radiation flux.

### 3 Radiative Transfer with Constraints; Escape Probability

#### 3.1 Two-level Atom

The simplest situation where we have a coupling of the radiative transfer equation and the statistical equilibrium equation is an idealized case of a two-level atom. Real atoms contain many energy levels, so that this approximation may seem at first sight to be grossly inadequate. However it actually provides a surprisingly good description of line formation in many cases of interest. And, more importantly, the case of the two level atom has a significant pedagogical value because it provides an explanation of many elementary processes that are crucial to understand NLTE line formation. In other words, a good physical understanding of line formation in a two-level atom is a prerequisite to understanding of more complicated cases. Therefore, this model will be discussed here in certain detail.

Let us first derive the expression for the source function. Figure 2 shows schematically the energy levels and all the elementary processes populating



**Fig. 2.** Schematic representation of microscopic processes in a two-level atom

and depopulating the levels. The absorption and emission coefficients are given by

$$\kappa_\nu = \frac{h\nu_0}{4\pi} (n_1 B_{12} - n_2 B_{21}) \phi(\nu) , \quad (72)$$

and

$$\eta_\nu = \frac{h\nu_0}{4\pi} n_2 A_{21} \phi(\nu) , \quad (73)$$

where  $\nu_0$  is the line-center frequency, and  $B_{12}$ ,  $B_{21}$  and  $A_{21}$  are the Einstein coefficients for absorption, stimulated emission, and spontaneous emission, respectively, for the radiative transitions between levels 1 and 2;  $n_1$  and  $n_2$  are populations (occupation numbers) of levels 1 and 2, respectively, and  $\phi(\nu)$  is the *absorption profile*. The latter expresses the probability density that if a photon is absorbed (emitted) in a line 1-2, it has a frequency in the range  $(\nu, \nu + d\nu)$ . The profile coefficient is thus normalized to unity,  $\int_0^\infty \phi(\nu) d\nu = 1$ . We assume that there is no other absorption or emission mechanism present.

It is advantageous to introduce a dimensionless frequency,  $x$ , by

$$x \equiv \frac{\nu - \nu_0}{\Delta\nu_D} , \quad (74)$$

$\Delta\nu_D$  is the Doppler width, given by  $\Delta\nu_D = (\nu_0/c)v_{\text{th}}$ , with the thermal velocity  $v_{\text{th}} = (2kT/m)^{1/2}$ ,  $m$  being the mass of the radiating atom. In the case of a pure Doppler profile (i.e. no intrinsic broadening of the spectral line; the only broadening is due to the thermal motion of radiators), the absorption profile is given by

$$\phi(x) = \exp(-x^2)/\sqrt{\pi} . \quad (75)$$

In a more general case, where there is an intrinsic broadening of lines described by a Lorentz profile in the atomic rest frame (the most common types of intrinsic broadening being the natural, Stark, and Van der Waals broadening - see Mihalas 1978, or monograph by Griem 1974), the profile function is given by the Voigt function,

$$\phi(x) = H(a, x)/\sqrt{\pi}, \quad H(a, x) = \frac{a}{\pi} \int_{-\infty}^{\infty} \frac{e^{-y^2}}{(x-y)^2 + a^2} dy . \quad (76)$$

The Voigt function is a convolution of the Doppler profile (i.e. the thermal motions) and the Lorentz profile (intrinsic broadening). The parameter  $a$  is a damping parameter expressed in units of Doppler width,  $a = \Gamma/(4\pi\Delta\nu_D)$ , where  $\Gamma$  is the atomic damping parameter. For instance, for the natural broadening of a line originating in a two-level atom,  $\Gamma = A_{21}$ .

Opacity in the line may be written as

$$\kappa_x = \kappa \phi(x) , \quad (77)$$

and analogously for  $\eta_x$ . The optical depth corresponding to the frequency-independent opacity,  $\kappa$ , is called the *frequency-averaged opacity* in the line, and is often used in line transfer studies. Notice that this opacity is *not* equal to the line center opacity,  $\kappa(0)$ , but is related to it by, for instance for the Doppler profile,  $\kappa(0) = \kappa/\sqrt{\pi}$ .

A remark is in order. We use the same profile coefficient for absorption, stimulated emission, and spontaneous emission - all of them are given through  $\phi(\nu)$ . This is an approximation called *complete redistribution* (CRD), which holds if an emitted photon is completely uncorrelated to a previously absorbed photon. In other words, the absorbed photon is re-emitted, i.e. redistributed, *completely*, without any memory of the frequency at which it was previously absorbed. A more exact description, taking into account photon correlations, is called the *partial redistribution* (PRD) approach. A discussion of this approach is beyond the scope of the present lecture; moreover, PRD effects are important only for certain lines (e.g. strong resonance lines, like hydrogen  $L\alpha$ , Mg II h and k lines, etc.), and under certain conditions (rather low density). The interested reader is referred to several reviews (e.g. Mihalas 1978; Hubeny 1985).

The source function follows from (72) and (73),

$$S_\nu \equiv \frac{\eta_\nu}{\kappa_\nu} = \frac{n_2 A_{21}}{n_1 B_{12} - n_2 B_{21}} \equiv S^L , \quad (78)$$

which is independent of frequency, thanks to the approximation of CRD.

Next step is to determine the ratio  $n_2/n_1$  which enters the source function. This is obtained from the statistical equilibrium equation, which states that the number of transitions into the state 1 (or 2) is equal to the number of transitions out of state 1 (2). This equation reads

$$n_1 (R_{12} + C_{12}) = n_2 (R_{21} + C_{21}) , \quad (79)$$

where  $R$ 's are the radiative rates, and  $C$ 's the collisional rates. The radiative rates are given by

$$R_{12} = B_{12} \int_0^\infty J_\nu \phi(\nu) d\nu \equiv B_{12} \bar{J} , \quad (80)$$

$$R_{21} = A_{21} + B_{21} \int_0^\infty J_\nu \phi(\nu) d\nu \equiv A_{21} + B_{21} \bar{J} , \quad (81)$$

where the quantity  $\bar{J}$  is called the *frequency-averaged mean intensity* of radiation. We will view here collisional rates as known functions of electron density (since collisions with electrons are usually most efficient) and temperature; for details, refer e.g. to Mihalas (1978).

Using the well-known relations between the Einstein coefficients,  $B_{21}/B_{12} = g_1/g_2$ , and  $A_{21}/B_{21} = 2h\nu_0^3/c^2$ , and the relation between the collisional rates,  $C_{21}/C_{12} = (n_1/n_2)^* = (g_1/g_2) \exp(h\nu_0/kT)$  (where  $(n_1/n_2)^*$  denotes the LTE population ratio), we obtain after some algebra

$$S = (1 - \epsilon) \bar{J} + \epsilon B_{\nu_0} , \quad (82)$$

where

$$\epsilon = \frac{\epsilon'}{1 + \epsilon'} ; \quad \epsilon' = \frac{C_{21}(1 - e^{-h\nu/kT})}{A_{21}} . \quad (83)$$

In the typical case,  $h\nu/kT \gg 1$  (since typical resonance lines, for which the two-level approximation is adequate, are formed in the UV region where the frequency is large), and therefore  $\epsilon$  may be expressed simply as

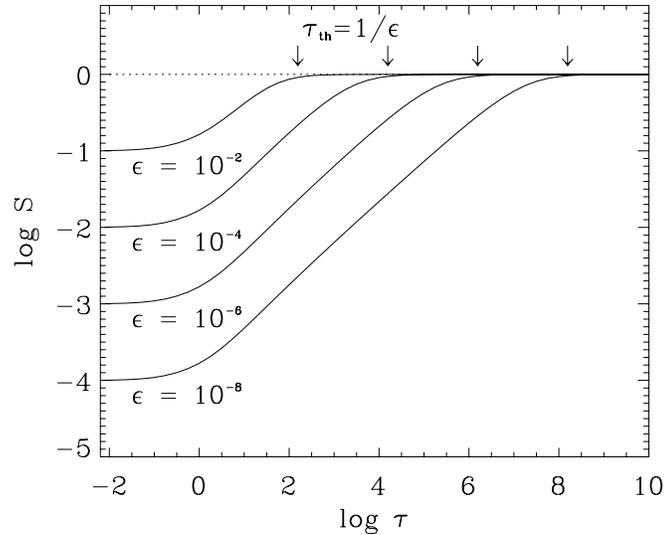
$$\epsilon \approx \frac{C_{21}}{C_{21} + A_{21}} , \quad (84)$$

which shows that  $\epsilon$  may be interpreted as a *destruction probability*, i.e. the probability that an absorbed photon is destroyed by a collisional de-excitation process ( $C_{21}$ ) rather than being re-emitted ( $A_{21}$ ).

Equation (82) is the fundamental equation of the problem. The first term on the right hand side represents the photons in the line created by scattering, i.e. by the emission following a previous absorption of a photon, while the second term represents the thermal creation of a photon, i.e. an emission following a previous collisional excitation.

Mathematically, the source function, (82), is still a *linear* function of the mean intensities. This is the case only for a two-level atom; in a general multi-level atom the source function contains non-linear terms in the radiation intensity. The two-level atom is thus an interesting pedagogical case: it contains a large-scale coupling of the radiation field and matter, yet the coupling, although being non-local, is still linear, and therefore much easier to handle (and understand!) than in the general case.

By applying any of the numerical methods which are discussed in the next chapter, one can easily obtain a solution of the two-level atom problem. Let us



**Fig. 3.** Source function for a two-level atom in a constant-property semi-infinite atmosphere, with  $B = 1$  (which only states that the source function is expressed in units of  $B$ ), and for various values of the destruction parameter  $\epsilon$ ;  $\epsilon = 10^{-2}, 10^{-4}, 10^{-6}, 10^{-8}$

take a standard example of line formation in a homogeneous semi-infinite slab. The homogeneity implies that all material properties (temperature, density, etc.) are independent of depth. In the context of the source function, (82), this means that  $\epsilon$ ,  $B$ , and  $\phi(x)$  are depth-independent. The solution, first obtained by Avrett and Hummer (1965), is displayed in Fig. 3 for several values of the destruction parameter  $\epsilon$ . It shows two interesting features:

i) The surface value of the source function is equal to  $\sqrt{\epsilon} B$ . Actually, this is a rather robust result, which is valid regardless of the type of the profile coefficient. Several rigorous mathematical proofs exist (see, e.g., monograph by Ivanov 1973); a physical explanation of this result was given by Hubeny (1987).

ii) The source function starts to deviate from the Planck function at a certain depth; below this point it is essentially equal to  $B$ . This depth is called the *thermalization depth*, and is traditionally denoted as  $A$ . We use here the notation  $\tau_{\text{th}}$  to avoid confusion with the  $A$ -operator. Figure 3 indicates that for a Doppler profile,  $\tau_{\text{th}} \approx 1/\epsilon$ . This indeed agrees with a more rigorous analytical study (Avrett and Hummer 1965; Ivanov 1973). These analyses

moreover show that for a Voigt profile the thermalization depth is even larger,  $\tau_{\text{th}} \approx a/\epsilon^2$ .

Why does the source function decrease towards the surface? We know that in the case of a homogeneous medium the departures from LTE arise only because of the presence of the boundary through which the photons escape. Before the line photons “feel” the presence of the boundary (i.e. in large enough optical depths), all microscopic processes depicted on Fig. 2 are in detailed balance, so the LTE approximation holds. However, as soon as the photons start to feel the boundary, i.e. they start to escape from the medium through the boundary, the photo-excitations are no longer balanced by radiative de-excitations. Since the absorption rate depend on the number of photons present, while the spontaneous emission rate does not (we neglect for simplicity the stimulated emission), the number of radiative excitations drops below the number of de-excitations as soon as photons start to escape. The lower level will consequently start to be overpopulated with respect to LTE, while the upper level will be underpopulated. Since the source function measures the number of photons created per unit optical depth, and since the number of created photons is proportional to the population of the upper level (because this is the level from which the atomic transition accompanied by the photon emission occur), the source function has to drop below the Planck function.

Having understood that, we now face an intriguing question: Given that departures from LTE arise because of the presence of the boundary, how come that the thermalization depth, i.e. the depth where the departures of the source function from the Planck function set in, is so large? Recall that the optical depth  $\tau$  in Fig. 3 is the frequency-averaged optical depth in the line. One might then expect that the presence of the boundary is felt by an “average” photon around  $\tau \approx 1$ , while the actual depth where photons feel the boundary is much larger (e.g.  $\tau \approx 10^6$  for a typical value of  $\epsilon = 10^{-6}$ )!

The explanation hinges on the fact that an “average” photon is not the one which is responsible for the transport and escape of photons in a line. Let us follow a photon trajectory from the point of its thermal creation. Let us assume that the photon was created at a large optical distance from the boundary. The photon is created with a large probability of having the frequency near the line center, because this probability is given by the absorption profile,  $\phi(x)$ , which is a sharply peaked function of frequency around  $x = 0$ . Consequently, the monochromatic optical depth is large, and so the physical distance it travels before the next absorption (i.e. the geometrical distance corresponding to  $\tau_x \approx 1$ ) is quite small. The same situation very likely occurs after the next scattering. We are then left with the following picture of photon trajectory in the two-level atom case with complete redistribution (the trajectory for the case of partial redistribution is quite different!): The photon makes many consecutive scatterings with the frequency staying close to the line center; during these scatterings the photon practically does not

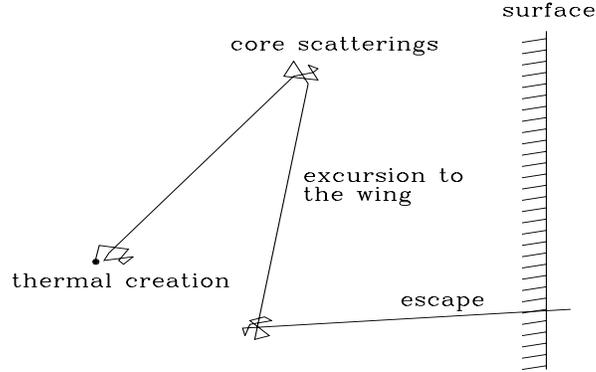


Fig. 4. Schematic representation of a trajectory of a photon in a gas of two-level atoms

move at all in the physical space. However, in a very infrequent event when it is re-emitted in the wing, the opacity it sees drops suddenly by orders of magnitude, and therefore it can travel a very large distance. The situation is depicted in Fig. 4. We see that the transfer in the core is inefficient; what really accomplishes the transfer are infrequent excursions of the photon to the line wings. This makes the photon transfer quite different from the massive particle transport: The particle mean free path remains of the same order of magnitude when a particle diffuses through the physical space, while the photon mean free path can change enormously. It is now clear why the thermalization depth is so large: it is determined by line-wing photons, whose mean free path is much larger than that of the core photons, which in turn define the mean optical depth  $\tau$ .

It is also clear why the thermalization depth depends on the destruction probability  $\epsilon$ . The total number of consecutive scatterings is of the order of  $1/\epsilon$ ; if the photon does not escape before it experiences  $1/\epsilon$  scatterings, it is destroyed by collisional processes, and therefore does not feel the presence of the boundary. These considerations are made more quantitative by the *escape probability* approach, which we shall consider in detail in the next subsection.

Finally, I mention that the source function for a line in a general multi-level atom can always be written in a form analogous to (82), viz. (see, e.g., Mihalas 1978)

$$S_{ij}^L = (1 - \epsilon_{ij}) \bar{J}_{ij} + \eta_{ij} , \quad (85)$$

where  $\epsilon_{ij}$  and  $\eta_{ij}$  are the generalized destruction and creation terms, re-

spectively; subscripts  $ij$  indicate that the quantities are appropriate for the transition  $i \rightarrow j$ . This approach is called the *equivalent-two-level-atom* (ETA) approach. The source function is formally a linear function of the mean intensity. However, it should be realized that the destruction and creation terms  $\epsilon_{ij}$  and  $\eta_{ij}$  contain contributions from the transition rates in all transitions in and out of states  $i$  and  $j$ , which depend on the radiation field. Therefore, despite apparent linearity of equation (85), one has to solve a general multi-level atom problem by an iteration process. The ETA approach may or may not converge in actual situations, and is not recommended as a robust and universal method. Nevertheless, it may be useful in some applications (see, for instance, several papers in Kalkofen 1984 and 1987; or Castor et al. 1992).

### 3.2 Escape Probability

Let us first consider a probability that a photon with frequency  $\nu$  and propagating in the direction specified by angle  $\mu$  escapes in a single flight. This probability is given by

$$p_{\nu\mu} = e^{-\tau_{\nu\mu}} \quad , \quad (86)$$

which follows from the very physical meaning of optical depth (see Sect. 2.4). The angle-averaged escape probability is given by

$$p_{\nu}(\tau_{\nu}) = \frac{1}{2} \int_0^1 e^{-\tau_{\nu}/\mu} d\mu = \frac{1}{2} E_2(\tau_{\nu}) \quad , \quad (87)$$

where the integration only extends for angles  $\mu \geq 0$ , since photons moving in the inward direction ( $\mu < 0$ ) cannot escape. Finally, the angle- and frequency-averaged escape probability for photons in one line is given by (adopting the  $x$ -notation, and writing  $x$  as a subscript)

$$p_e(\tau) = \int_{-\infty}^{\infty} \phi_x p_x(\tau_x) dx = \frac{1}{2} \int_{-\infty}^{\infty} \phi_x E_2(\tau \phi_x) dx \quad . \quad (88)$$

Notice that at the surface,  $p_e(0) = 1/2$ , because a photon is either emitted in the outward direction, in which case it certainly escapes, or in the inward direction, in which case it does not escape (assuming an isotropic emission).

We may now quantify the considerations given in the previous subsection. We introduce the *photon destruction probability* by

$$p_d = \epsilon \quad , \quad (89)$$

and we have the *photon escape probability*,  $p_e$ , defined above. Now, if  $p_e \ll p_d$ , photons are likely thermalized before escaping from the medium. In other words, the line photons do not feel the presence of the boundary, and therefore  $S \approx B$ . On the other hand, if  $p_e \gg p_d$ , photons likely escape before being

thermalized, i.e. destroyed by a collisional process. It is therefore natural to define the thermalization depth  $\tau_{\text{th}}$ , as

$$p_e(\tau_{\text{th}}) = p_d \quad , \quad (90)$$

which indeed gives, by substituting the Doppler profile in (88), the expression  $\tau_{\text{th}} \approx 1/\epsilon$ .

The escape probability considerations are actually much more powerful than just to explain the value of thermalization depth. One may in fact construct approximate expressions for the source function as a function of depth. To demonstrate this, let us consider the following simple model: We know that  $\bar{J}$  measures the number of photons absorbed in a line per unit optical depth interval [which may be verified by integrating (35) over frequencies]. If we are far from the surface, all the photons emitted per unit optical distance,  $S(\tau) d\tau$ , either escape from the medium by a single flight (with a probability  $p_e$ ), or are re-absorbed, more or less on the spot, with probability  $1 - p_e$ . This suggests that the number of photons absorbed at  $\tau$ , i.e.  $\bar{J}(\tau)$ , should be given by

$$\bar{J}(\tau) = S(\tau) (1 - p_e) \quad , \quad (91)$$

which gives us the desired approximate relation between the averaged mean intensity of radiation and the source function, without actually solving the transfer equation!

A very interesting point is that we can arrive, purely mathematically, to the same equation if we start with the integral expression (116) (see Sect. 4.2), and do the following trick: Since the kernel function  $K_1(t)$  varies much more rapidly than  $S(t)$ , we may assume that the source function does not vary over the range where the kernel function varies appreciably. In other words, we may remove  $S(t)$  from the integral in (116), and put  $S(t) = S(\tau)$ . One may easily verify that by integrating the kernel function  $K_1$  over  $\tau$  one obtains (91) with  $p_e$  given by (88).

Substituting (91) into the the expression for the source function, (82), we obtain the following expression for the source function,

$$S(\tau) = \frac{\epsilon}{\epsilon + (1 - \epsilon)p_e} B \quad , \quad (92)$$

which is traditionally called the *first-order escape probability approximation*. It describes very well the behavior of the source function at depths, but it fails to reproduce the  $\sqrt{\epsilon}$ -law, since it yields for the source function at the surface  $S(0) = 2\epsilon/(1 + \epsilon) B$ , which may be quite different from  $\sqrt{\epsilon} B$ . The reason for this can be easily understood: any transfer of photons is neglected here, and the problem is reduced to just two mechanisms – a photon either escapes in a single direct flight, or is thermalized. This so-called “dichotomous” model works well deep in the atmosphere, but fails in the outer layers of the atmosphere, where the transfer of photons is important.

Without going to any more details, I just mention that the so-called *second-order escape probability* formalism, which takes into account some aspects of the photon transport, was developed (for an illuminating discussion, see an excellent review by Rybicki, 1984). The resulting expression for the source function in a homogeneous atmosphere is

$$S(\tau) = \left( \frac{\epsilon}{\epsilon + 2(1 - \epsilon)p_e} \right)^{1/2} B , \quad (93)$$

which behaves very similarly to the first-order approximation at depths, but now yields the correct expression for the source function at the surface,  $S(0) = \sqrt{\epsilon} B$ .

Concluding, the escape probability approach is very useful and very powerful, because it is able to provide simple approximate relations between the source function and the mean intensity of radiation, based on simple physical arguments. It can therefore be used in cases where detailed numerical solutions are either too complicated and time consuming (like in the case of radiation hydrodynamic simulations, where the radiative transfer equation is solved in a huge number of time steps), or where a high accuracy of predicted emergent radiation is not required. However, one should always keep in mind that the escape probability methods are *inherently approximate*, and therefore one should be always aware of their potential limitations and inaccuracies. Finally, I stress that these methods were discussed here partly because of the above reasons, and partly because of their intimate relation to a class of modern numerical methods, called Accelerated Lambda Iteration (ALI) methods, which will be discussed in the next section.

## 4 Numerical Methods

There are several types of numerical method, depending on the degree of complexity of the problem at hand. In this section, we will consider numerical methods for treating three basic problems, ordered by increasing complexity, i) a formal solution of the radiative transfer equation – where the source function is specified; ii) a solution of linear line formation problems – the source function is a linear function of radiation intensity; and iii) a solution of general non-linear problems.

### 4.1 Formal Solution of the Transfer Equation

By the term formal solution we understand a solution of the transfer equation if the source function is fully specified. We have already shown the formal solution of the transfer equation, given by (39) for the general case; or by (55) and (56) for a semi-infinite atmosphere. The related expression for the mean intensity is (57). In practice, we may replace the integral over optical

depth by a quadrature sum, and calculate the radiation intensity by a simple summation.

Why, then, would we need to consider other numerical methods for this apparently trivial problem? The basic point is that the simple numerical quadrature is extremely inefficient from the point of view of computer time. This is because the kernel functions contain exponentials, which are very costly to compute. As we will see later on, the speed of modern numerical methods which solve a general coupled problem is in fact determined by the speed with which the individual formal solutions are accomplished. Therefore, we have to seek as efficient numerical schemes for performing a formal solution as possible.

There are essentially two classes of methods, namely those based on

1. the first-order form of the transfer equation; or
2. the second-order form of the transfer equation. The second-order method is usually called the Feautrier method, in honor of its originator (Feautrier 1964).

**First-order methods.** They were not used very much during the last two decades. However, they were revived recently by an ingenious adaptation of the Discontinuous Finite Element (DFE) method by Castor et al. (1992). This scheme now appears to be an extremely advantageous method, and will very likely be used more and more in the stellar atmosphere numerical work. I will present only a brief outline here; the interested reader is referred to the original paper.

Let us assume a given frequency  $\nu$  and angle  $\mu$ . Let us denote  $\tau$  the monochromatic optical depth at frequency  $\nu$ , along the ray specified by angle  $\mu$ . In the following, we drop an explicit indication of frequency and angle variables. The intensity of radiation in the optical depth interval between two discretized depth points,  $(\tau_d, \tau_{d+1})$ , is assumed to be given as a linear function of optical depth.

$$I(\tau) = I_d^+ \frac{\tau_{d+1} - \tau}{\Delta\tau_d} + I_{d+1}^- \frac{\tau - \tau_d}{\Delta\tau_d} . \quad (94)$$

To avoid confusion, I stress that we deal with the intensity in *one direction* only; the notation  $I^+$  and  $I^-$  does not mean intensities in opposite directions, as it is usually used in the radiative transfer theory.

If  $I_d^- = I_d^+$ , the linear representation of intensity, (94), is a continuous function of frequency. However, the related numerical method would be quite inaccurate. The essence of the DFE method is to allow for step *discontinuities* at points  $\tau_d$ , i.e. we consider generally  $I_d^- \neq I_d^+$ . Substituting (94) into the transfer equation (29), and performing analytic manipulations described in Castor et al. (1992), one obtains final linear relations for the quantities  $I^+$  and  $I^-$ , viz.

$$\frac{I_{d+1}^- + I_d^+ - 2I_d^-}{\Delta\tau_d} = S_d - I_d^- , \quad (95)$$

and

$$\frac{I_{d+1}^- - I_d^+}{\Delta\tau_d} = S_{d+1} - I_{d+1}^- . \quad (96)$$

By eliminating  $I_d^+$  we obtain a simple linear recurrence relation for  $I_d^-$ ,

$$(\Delta\tau_d^2 + 2\Delta\tau_d + 2)I_{d+1}^- - 2I_d^- = \Delta\tau_d S_d + \Delta\tau_d(\Delta\tau_d + 1)S_{d+1} , \quad (97)$$

and  $I_d^+$  follows from

$$(\Delta\tau_d^2 + 2\Delta\tau_d + 2)I_d^+ = 2(\Delta\tau_d + 1)I_d^- + \Delta\tau_d(\Delta\tau_d + 1)S_d - \Delta\tau_d S_{d+1} . \quad (98)$$

Finally, the resulting specific intensity at  $\tau_d$  is given as a linear combination of the “discontinuous” intensities  $I_d^-$  and  $I_d^+$ ,

$$I_d = \frac{I_d^- \Delta\tau_d + I_d^+ \Delta\tau_{d-1}}{\Delta\tau_d + \Delta\tau_{d-1}} . \quad (99)$$

**Second-order, or Feautrier method.** The basis of the method is to introduce the symmetric and antisymmetric averages of the specific intensity,

$$j_{\mu\nu} \equiv \frac{1}{2}[I(+\mu, \nu) + I(-\mu, \nu)] , \quad (0 \leq \mu \leq 1) , \quad (100)$$

$$h_{\mu\nu} \equiv \frac{1}{2}[I(+\mu, \nu) - I(-\mu, \nu)] , \quad (0 \leq \mu \leq 1) . \quad (101)$$

Considering separately the transfer equation (29) for positive and negative  $\mu$ 's, and adding and subtracting these equations we obtain,

$$\mu(dh_{\mu\nu}/d\tau_\nu) = j_{\mu\nu} - S_\nu , \quad (102)$$

$$\mu(dj_{\mu\nu}/d\tau_\nu) = h_{\mu\nu} . \quad (103)$$

Using (103) to eliminate  $h_{\mu\nu}$  from (102), we obtain

$$\mu^2 \frac{d^2 j_{\mu\nu}}{d\tau_\nu^2} = j_{\mu\nu} - S_\nu . \quad (104)$$

This equation is very similar to the moment equation (54); also the quantity  $j_{\mu\nu}$  is very similar to the mean intensity  $J_\nu$ . The essential difference between (54) and (104) is that (104) is a *closed* equation for the symmetrized intensity  $j_{\mu\nu}$ , which may therefore be solved in a single step if the source function is known.

Special care should be devoted to the boundary conditions. The specific intensity is specified for negative  $\mu$ 's at the upper boundary, and for the positive  $\mu$ 's at the lower boundary;

$$I(-\mu, \nu, \tau = 0) = I_{\mu\nu}^- , \quad (0 \leq \mu \leq 1) , \quad (105)$$

$$I(+\mu, \nu, \tau = \tau_{\max}) = I_{\mu\nu}^+ , \quad (0 \leq \mu \leq 1) , \quad (106)$$

( $\tau_{\max} = \infty$  for a semi-infinite atmosphere). Substituting (105) into (103), and using (101), we obtain

$$\mu (dj_{\mu\nu}/d\tau_\nu)_0 = j_{\mu\nu}(0) - I_{\mu\nu}^- , \quad (107)$$

$$\mu (dj_{\mu\nu}/d\tau_\nu)_{\tau_{\max}} = I_{\mu\nu}^+ - j_{\mu\nu}(\tau_{\max}) . \quad (108)$$

In most cases, the incoming intensity  $I_{\mu\nu}^- = 0$ . For a semi-infinite atmosphere, the diffusion approximation is usually used for the lower boundary condition,

$$I_{\mu\nu}^+ = B_\nu(\tau_{\max}) + \mu(\partial B_\nu/\partial\tau_\nu)_{\tau_{\max}} . \quad (109)$$

Equation (104), together with boundary conditions (107) and (108) is solved numerically by discretizing the depth variable. The discretized form may be written as (writing  $u \equiv j_{\mu\nu}$ ),

$$-A_d u_{d-1} + B_d u_d - C_d u_{d+1} = S_d , \quad (110)$$

Detailed expressions for the elements  $A, B, C$  are given in the standard textbooks (e.g. Mihalas 1978). The resulting tridiagonal set of equations is solved by a straightforward Gaussian elimination, consisting in a forward-backward recursive sweep, namely

$$D_d = (B_d - A_d D_{d-1})^{-1} C_d, \quad D_1 = B_1^{-1} C_1 , \quad (111)$$

$$Z_d = (B_d - A_d D_{d-1})^{-1} (S_d + A_d Z_{d-1}), \quad Z_1 = B_1^{-1} S_1 , \quad (112)$$

followed by the reverse sweep,

$$u_d = D_d u_{d+1} + Z_d, \quad u_{ND+1} = 0 , \quad (113)$$

where  $ND$  is the number of discretized depth points.

## 4.2 Linear Coupling Problems

The second class of methods are those in which the source function is given as a known, *linear*, function of the specific intensities. A typical example is the line formation in a two-level atom, where the source function is given by (82). A more general case is the equivalent-two-level-atom source function, (85), with the creation and destruction terms  $\eta_{ij}$  and  $\epsilon_{ij}$  assumed to be specified. In other words, this corresponds to solving the transfer problem for one line at a time.

Numerical solution can either be done by a *differential equation* approach, or by an *integral equation* approach.

**The differential equation method** consists in choosing discrete values of frequencies,  $(x_i, i = 1, \dots, NF)$  and angles  $(\mu_j, j = 1, \dots, NA)$ , and to solve a coupled set of transfer equations written for all frequency-angle points,

$$\mu_j \frac{dI(x_i, \mu_j, \tau)}{d\tau} = \phi(x_i) [I(x_i, \mu_j, \tau) - S(\tau)] \quad , \quad (114)$$

where the source function on the right hand side is given by (85), replacing the integrals over frequency and angle by a quadrature sum,

$$S = (1 - \epsilon) \frac{1}{2} \sum_{i=1}^{NF} \sum_{j=1}^{NA} w_i^x w_j^\mu \phi(x_i) I(x_i, \mu_j) + \eta \quad , \quad (115)$$

where  $w_i^x$  and  $w_j^\mu$  are the quadrature weights for the integration over frequencies and angles, respectively. The source function couples *all* frequencies and angles, but the main point is that the source function is a linear function of the specific intensities, (85). The system (114) is thus a system of *linear* differential equations. One may construct a column vector  $\mathbf{I}$  whose elements are values of specific intensity at given depth for all pairs of  $(x, \mu)$ , and write all equations (114) as one differential equation for the vector  $\mathbf{I}$ . One may then apply the Feautrier method described above; equations (100) - (113) remain the same, only the meaning of  $u$  and the coefficients  $A, B, C$  will be different.  $u$  will represent a vector  $(j_{\nu_i \mu_j}, i = 1, \dots, NF, j = 1, \dots, NA)$  (i.e. the Feautrier intensities at all discretized frequency-angle points), and  $A, B, C$  will be  $(NF \times NA) \times (NF \times NA)$  matrices. The resulting system of linear equations forms a block-tridiagonal system.

**The integral equation method** is based on expressing the averaged mean intensity  $\bar{J}$  as an integral over  $S$ , which easily follows from the formal solution of the transfer equation discussed in Sect. 2. By integrating (57) over frequencies, we obtain

$$\bar{J}(\tau) = \int_0^\infty S(t) K_1(|t - \tau|) dt \quad , \quad (116)$$

where the kernel function  $K_1$  is given by

$$K_1(s) = \int_0^\infty E_1(\phi_x s) \phi_x^2 dx \quad . \quad (117)$$

The behavior of the kernel function depends on the type of the absorption profile. As can be intuitively expected, it has a narrower peak for the Doppler profile than for the Voigt profile. A useful numerical algorithm for computing the function  $K_1$  was given by Hummer (1981).

Substituting (116) into (82) yields the following integral equation for the source function:

$$S(\tau) = (1 - \epsilon) \int_0^\infty S(t) K_1(|t - \tau|) dt + \epsilon B \quad . \quad (118)$$

This equation was first solved more than three decades ago by Avrett and Hummer (1965). The equation was subsequently extensively studied analytically by the Russian analytical school. Many elegant analytical results are summarized in a monograph by Ivanov (1973); this book is recommended to anyone who intends to study the radiative transfer seriously.

The integral equation approach has several advantages and drawbacks. The advantage is that it deals with one simple integral equation for  $S$ , so in a sense it is formulated in the most efficient way since the knowledge of  $S$  represents the solution of the problem (individual specific intensities of radiation may then easily be obtained by the formal solution of the transfer equation). In other words, the coupling of radiation and material properties in the integral equation approach is fully contained in the function  $K_1$ , which is calculated in advance, while in the differential equation approach the coupling is treated explicitly. Nevertheless, the differential equation approach may be reformulated in an efficient way by casting it in the form analogous to the integral approach (the so-called Rybicki variant of the Feautrier method – see Rybicki 1971). In any case, the integral equation approach suffers from a significant drawback, namely that in evaluating the kernel function (and in the formal solution of the transfer equation), one faces the task of evaluating a large number of exponentials, which are computationally very costly. Therefore, most of the actual numerical work in the radiative transfer is nowadays being done using the differential equation approach.

### 4.3 Accelerated Lambda Iteration

In the previous section, we saw that the two-level atom problem is a linear one, and thus may be solved in a single step, without any iterations. However, one pays a high price for that: one has to invert, at every discretized depth point, some auxiliary matrices whose dimension is given by the product of the number of discretized frequencies times the the number of discretized angles. [The situation may be alleviated by employing the so-called Variable Eddington Factor technique, developed by Auer and Mihalas (1970), which treats the angle coupling separately. The size of matrices is reduced but is still given by the number of frequencies, which may be large]. Generally, one should realize that *any method that describes a coupling of various quantities by means of a direct matrix inversion is fundamentally limited* in that the computer time scales as the cube of the number of quantities (i.e. the number of frequency points in our case).

Therefore, one needs faster schemes. How can this be accomplished? The clue is to realize that some part of the physical coupling is more important than others. In other words, not all the parts of the coupling should necessarily be treated on the same footing; it is more or less a numerical overkill to do so. So, this hints that the “important part” of the coupling should be treated exactly, while the rest may be treated iteratively.

Below, I demonstrate the method on an example of a two-level atom. However, the method is much more powerful, and can be applied to virtually any astrophysical radiative transfer problem. One such application will be mentioned in Sect. 5.4. We first recall that the two-level problem may be written, by substituting (60) into (82), as

$$S = (1 - \epsilon)\bar{A}[S] + \epsilon B \quad , \quad (119)$$

(which is just another expression of the integral equation form 118). The frequency-averaged lambda operator is given by

$$\bar{A} = \int_0^\infty A_\nu \phi(\nu) d\nu \quad , \quad (120)$$

with the frequency-dependent Lambda operator  $A_\nu$  given by (60). In the following, I omit the bar over  $A$  for notational simplicity.

In a seminal paper Cannon (1973) introduced into astrophysical radiative transfer theory the *method of deferred corrections* (also called, somewhat inaccurately, an *operator splitting*), long known in numerical analysis. The idea consists of writing

$$A = A^* + (A - A^*) \quad , \quad (121)$$

where  $A^*$  is an appropriately chosen *approximate lambda operator*. The iteration scheme for solving (119) may then be written as

$$S^{(n+1)} = (1 - \epsilon)A^*[S^{(n+1)}] + (1 - \epsilon)(A - A^*)[S^{(n)}] + \epsilon B \quad , \quad (122)$$

or, in a slightly different form whose importance becomes apparent later,

$$S^{(n+1)} - S^{(n)} = [1 - (1 - \epsilon)A^*]^{-1}[S^{\text{FS}} - S^{(n)}] \quad , \quad (123)$$

where

$$S^{\text{FS}} = (1 - \epsilon)A[S^{(n)}] + \epsilon B \quad . \quad (124)$$

Superscript FS stands for Formal Solution. In other words, (122) shows that the action of the exact  $A$  operator is split into two contributions: an approximate  $A^*$  operator which acts on the new iterate of the source function, and the difference between the exact and approximate operator,  $A - A^*$ , acting on the previous, known, iterate of the source function. The latter contribution may be easily evaluated by the formal solution.

If we choose  $A^* = 0$ , we recover the “ordinary” lambda iteration, which is straightforward, but is known to converge very slowly – see, e.g. Mihalas 1978; or Olson, Auer, Buchler (1986 – hereafter referred to as OAB). On the other hand, the choice  $A^* = A$  represents the exact method, which is done without any iteration, but an inversion of the exact  $A$  operator may be costly. So, in order that  $A^*$  brings an essential improvement over both methods, it has to satisfy the following requirements: i) it has to incorporate all the *essential* properties of the exact  $A$  operator in order to obtain a fast convergence rate of the iteration process; but at the same time, ii) it must be

easy (and cheap) to invert. These requirements are generally incompatible, therefore the construction of the optimum  $A^*$  is a delicate matter.

The interesting history of the quest for the optimum  $A^*$  operator is summarized by Hubeny (1992). Let us stress that a numerically most advantageous approximate operator is a diagonal (i.e. local) operator  $A^*$ , in which case it represents a multiplication by a scalar value, and its inversion is a simple division. To understand that the term “diagonal operator” is equivalent to the term “local operator”, recall (61) and (62). These equations also explain why a good approximation for the exact  $A$ -matrix is its diagonal. Recall that the matrix element  $A_{ji}$  tells us what portion of photons created in an elementary interval around depth point  $i$  [i.e.  $S(\tau_i)$ ] are being absorbed at depth point  $j$  [described by  $J(\tau_j)$ ]. Most photons are absorbed very close to the point of their creation, so the diagonal term  $A_{ii}$  is much larger than the off-diagonal terms. In other words, approximating the exact  $A$  by a diagonal operator means replacing the kernel function for the mean intensity, (60) by a  $\delta$ -function, which, as seen in Fig. 1, is quite reasonable. (These considerations also show that the next simplest approximation for the  $A$ -operator would be its tridiagonal part; here an interaction between a given depth and its immediate neighbors is taken into account.)

Equation (123) is particularly instructive. It shows that iteration is driven, similarly as the ordinary lambda iteration, by the difference between the old source function and the newer source function obtained by formal solution. However, unlike the ordinary lambda iteration, this difference is amplified by the “acceleration operator”  $[1 - (1 - \epsilon)A^*]^{-1}$ . To gain more insight, let us consider a diagonal (i.e. local)  $A^*$  operator. The appropriate  $A^*$  has to be chosen such as  $A^*(\tau) \rightarrow 1$  for large  $\tau$  (see below). Since in typical cases  $\epsilon \ll 1$ , the acceleration operator indeed acts as a large amplification factor. This interpretation was first introduced by Hamann (1985), who also coined the term “Accelerated Lambda Iteration” (ALI). The acronym ALI is also sometimes understood to mean “Approximate Lambda Iteration”. Other terms for ALI are *Operator Perturbation* (Kalkofen 1987), or *Approximate-Operator Iteration* (AOI; Castor et al. 1991, 1992). Finally, the term “accelerated  $A$  iteration” should not be confused with “acceleration of convergence”, discussed later on.

How do we know that  $A^*$  should approach unity at large depths? Here comes the intimate relation between the escape probability and the ALI methods, mentioned in Sect. 3.2. Recall that the escape probability formalism gives a relation between between the mean intensity and the source function, namely  $\bar{J} = (1 - p_e)S$ . This is exactly what we need here – a local approximate relation between  $\bar{J}$  and  $S$ . We may thus put, as a reasonable choice,  $A^* = 1 - p_e$ , which indeed shows that  $A^*$  approaches unity for large  $\tau$ . This escape-probability form of  $A^*$  may be used for numerical work, but modern approaches provided more efficient and robust ways to construct the approximate  $A^*$  operator.

I will not discuss here all possible variants of the  $\Lambda^*$  operator; the interested reader is referred to Hubeny (1992). I will only mention several important papers. First, Scharmer (1981) revived Cannon's original ideas, and constructed an ingenious  $\Lambda^*$  operator based on the Eddington-Barbier relation. Next, OAB have shown, using rigorous mathematics, that a nearly optimum  $\Lambda^*$  operator is a diagonal part of the true  $\Lambda$  operator. Olson and Kunasz (1987) showed that the tridiagonal and possibly higher multi-band parts of the lambda operator yield even more rapid convergence. Finally, Rybicki and Hummer (1991) used a formalism based on the Feautrier scheme, employing a very efficient algorithm for inverting a tridiagonal matrix, and demonstrated that the entire set of the diagonal elements of  $\Lambda$  can be found with an order of  $ND$  operations. This feature makes it the method of choice, since it avoids computing costly exponentials, a problem inherent to both previous approaches (OAB; Olson and Kunasz 1987).

**Acceleration of Convergence.** This is a highly technical topic, but is mentioned here because it has recently become an important ingredient of the ALI methods. Only a brief summary of the basic ideas is presented here. Any iterative scheme can be written in the form

$$\mathbf{x}^{(n+1)} = \mathbf{F} \cdot \mathbf{x}^{(n)} + \mathbf{x}^{(0)} , \quad (125)$$

where  $\mathbf{F}$  is called the *amplification matrix*. In the case of the linear transfer problem, (122), we have  $\mathbf{F}$  given by  $\mathbf{F} = [1 - (1 - \epsilon)\Lambda^*]^{-1} [(1 - \epsilon)(\Lambda - \Lambda^*)]$ , where  $\mathbf{x}^{(n)}$  is an  $n$ -th iterate of the source function.

As it is well known from linear algebra, any iteration method where the  $(n + 1)$ -th iterate is solely evaluated by means of the previous one converges only *linearly*. However, taking into account information from the earlier iterates, one may find faster schemes. I will not discuss these methods in any detail here, the interested reader is referred to the review papers by Auer (1987, 1991), or to the original papers cited therein. I just briefly mention that for the most popular scheme, the Ng acceleration, the general expression for the accelerated estimate of the solution in the  $n$ -th iteration is written

$$\mathbf{x}^{\text{acc}} = \left( 1 - \sum_{m=1}^M \alpha_m \right) \mathbf{x}^{(n)} + \sum_{m=1}^M \alpha_m \mathbf{x}^{(n-m)} , \quad (126)$$

where the coefficients  $\alpha$  are determined by a residual minimization. Practical expressions are given by OAB, Auer (1987, 1991), or Hubeny and Lanz (1992).

#### 4.4 Non-linear Coupling Problems

To illustrate the basic problem of applying ALI in multilevel problems, let us first write down the expression for the radiative rates. For simplicity, let us

consider only lines; the treatment of continua is analogous. The net transition rate for any line  $i \rightarrow j$  ( $i$  and  $j \neq i$  represent any states of an atom), is

$$R_{ji}^{\text{net}} = n_j A_{ji} - (n_i B_{ij} - n_j B_{ji}) \bar{J}_{ij} \quad , \quad (127)$$

The basic ALI equation, (121), gives for  $\bar{J}_{ij}$

$$\bar{J}_{ij} = \Lambda^* [S^{\text{new}}] + (\Lambda - \Lambda^*) [S^{\text{old}}] \quad . \quad (128)$$

Here the second term, which may be written as  $\Delta \bar{J}_{ij}^{\text{old}}$ , is known from the previous iteration. However, the first term contains  $S^{\text{new}}$  which is a complicated, and generally non-linear function of the “new” populations.

This is an unfortunate situation. By applying the ALI idea, we have succeeded to eliminate the radiation intensity from the rate equations, but at the expense of ending with a set of non-linear equations for the populations. We cope with this problem by one of the possible two ways:

1. *Linearization.* The usual way of solving the set of non-linear equations is by applying the Newton–Raphson method. This may be rather time consuming because each iteration requires to set up and to invert the Jacobi matrix of the system.
2. *Preconditioning.* This is an ingenious way to analytically remove inactive (scattering) parts of radiative rates from the rate equations, and to recover a linearity of the ALI form of the rate equations.

Let us demonstrate the idea of preconditioning on a simple case, where the total source function is given by the line source function  $S_{ij} = n_j A_{ji} / (n_i B_{ij} - n_j B_{ji})$  (i.e., the case of non-overlapping lines and no background continuum). Let us further assume that we have a local (diagonal) approximate  $\Lambda^*$  operator ( $\Lambda^*$  is then a real number). The net rate (127) may be written, after some algebra,

$$R_{ji}^{\text{net}} = n_j A_{ji} (1 - \Lambda_{ji}^*) - (n_i B_{ij} - n_j B_{ji}) \Delta \bar{J}_{ij}^{\text{old}} \quad , \quad (129)$$

which is indeed *linear* in the populations!

This is a very interesting expression. Notice first that the original net rate, (127), is represented by a subtraction of two large contributions, *all* emission minus all absorptions, while the result, the net rate, is rather small. Physically, this follows from the fact that most emissions (i.e. radiative transitions  $j \rightarrow i$ ) are those which immediately follow a previous absorption of a photon (transitions  $i \rightarrow j$ ), i.e. they are the part of a *scattering* process. In order to improve the numerical conditioning of the system of rate equations, we have to somehow eliminate the scattering contributions, i.e. to “precondition” the rates. An illuminating discussion of this topic is presented by Rybicki (1984).

In the ALI form of the net rate, (129), we see that deep in the atmosphere,  $\Lambda^* \rightarrow 1$ , so that the first term is indeed very small. Similarly, the second term is also small because  $\Delta \bar{J}_{ij}$  is small. In other words, the radiative rates are

indeed preconditioned. In the context of the ALI approach, this idea was first used by Werner and Husfeld (1985); a systematic study was presented by Rybicki and Hummer (1991, 1992), who have extended it to the case of general overlap of lines and continua.

## 5 Model Atmospheres

### 5.1 Definition and Terminology

By the term *model atmosphere* we understand a specification of all the atmospheric state parameters as functions of depth. Since the problem is very complex, we cannot construct analytic solutions. Therefore, we discretize the depth coordinate and consider a finite number of depth points – this number is typically of the order of several tens to few hundreds. A model atmosphere is then a table of values of the state parameters in these discretized depth points.

Which are the parameters that describe the physical state of the atmosphere? The list of parameters depend on the type of the model, i.e. on the basic assumptions under which the model is constructed. Traditionally, the list of state parameters includes only massive particle state parameters (e.g. temperature, density, etc.), but not the radiation field parameters. This might seem to be in sharp contrast of what was being stated before, namely that radiation intensity is in fact a crucial parameter. It indeed is, and in fact the radiation intensity is an important state parameter in the process of constructing the atmospheric structure. But, when the system of all structural equations, which includes the radiative transfer equation, is solved, we do not have to keep the radiation intensity in the list of state parameters which has to be stored in the table representing the model. The point is that once all the necessary material properties are given, we may easily determine the radiation field by a formal solution of the transfer equation.

The terminology is sometimes ambiguous. Some astronomers, mostly observers, understand by the term “model stellar atmosphere” a table of emergent radiation flux as a function of wavelength. This is understandable, since for many purposes the predicted radiation from a star is the only interesting information coming out of the model. Let us take an example of a widely used Kurucz (1979, 1994) grid of model atmospheres. For each combination of input stellar parameters ( $T_{\text{eff}}$ ,  $\log g$ , and metallicity), he publishes two tables; one is the “model atmosphere” in our definition, i.e. a relatively short table of values of temperature, electron density, etc., in all depth points; the second table is a table of emergent flux versus wavelength. In fact, many if not most workers use only this second table. A drawback of using the tabulated model flux is that it has a fixed wavelength resolution (in the case of Kurucz models, it is relatively coarse – 10 Å), and thus cannot be used for purposes which require a high-resolution predicted spectrum. On the other

hand, from the genuine model, one may easily compute a spectrum of any resolution.

Below, I summarize the basic types of model stellar atmospheres.

**I) Static models.** These are models constructed under the assumption of hydrostatic equilibrium. Consequently, these models apply only to atmospheric layers that are indeed close to hydrostatic equilibrium, i.e. the macroscopic velocity is small compared to the thermal velocity of atoms. These layers are traditionally called *stellar photospheres*. Basic input parameters are the effective temperature,  $T_{\text{eff}}$ , the surface gravity,  $g$  (usually expressed as  $\log g$ ), and chemical composition. Strictly speaking, one should give the values of abundances of all individual chemical species. In reality, one usually considers solar abundances, or some ratio of some or all abundances with respect to the solar one. If all elements but hydrogen and helium share the common abundance ratio with respect to the solar abundances, this ratio is called *metallicity*. There are some additional input parameters, like the microturbulent velocity, or, in the case of convective models, the mixing length (or some other parameters approximating the convection)

There are several basic types of models:

- LTE grey models. They are the simplest possible models, based on the assumption that the opacity is independent of frequency. They are not used any longer for spectroscopic work, but they are useful for providing an initial estimate in any iterative method for constructing more realistic models, and they are very useful for pedagogical purposes. For this reason, they will be discussed at length in the next section.
- LTE models. They are based on the assumption of LTE (see Sect. 1.3). Two state parameters, for instance temperature,  $T$ , and density,  $\rho$ , (or electron density,  $n_e$ ), suffice to describe the physical state of the atmosphere at any given depth.
- NLTE models. This is a rather ambiguous term which encompasses any model which takes into account some kind of a departure from LTE. In early NLTE models, the populations of only few of the low-lying energy levels of the most abundant species, like H and He, were allowed to depart from LTE; the rest was treated in LTE. There are two basic kinds of NLTE models, or approaches to include NLTE effects:
  - Models solving for the full structure. The codes of general use include an early H-He model atmosphere code described by Mihalas et al (1975), the Kiel code (Werner 1987); PAM (Anderson 1987), and a universal code TLUSTY (Hubeny 1988).
  - NLTE line formation (also called a restricted NLTE problem). Here, the atmospheric structure (temperature, density, etc.) is assumed to be known from previous calculations (either LTE or simplified NLTE), and is kept fixed, while only radiative transfer and statistical equilibrium for a chosen atom/ion is solved simultaneously. The

popular codes of this sort include DETAIL/SURFACE (Butler and Giddings 1985); MULTI (Carlsson 1986), and MALI (Rybicki and Hummer 1991).

- NLTE line-blanketed models. This is in fact a subset of the previous item. I consider it separately because these models represent a qualitatively new step in the model construction. They are models where NLTE is considered for practically *all* energy levels and transitions between them – lines and continua – that influence the atmospheric structure. The number of such lines may actually go to millions, so the problem is presently extremely demanding on the computer resources and ingenuity of the numerical methods used. In these models, it is no longer necessary to compute the atmospheric structure using simple atomic models, and recalculate NLTE line formation in individual atoms separately. These models will be discussed in more detail in Sect. 5.5.

**II) Unified models.** By definition, unified model atmosphere are those which relax the a priori assumption of hydrostatic equilibrium, and which thus treat the whole atmosphere ranging from an essentially static photosphere to a highly dynamical wind on the same footing. Ideally, this would mean solving self-consistently the set of hydrodynamic equations (2) - (4) and the radiative transfer equation. This is a tremendous task, which has not yet been even attempted to solve generally. Instead, one treats the hydrodynamic of the wind taking into account radiation in some approximate way (for instance, the line driven wind theory by Castor, Abbott, Klein 1975; or Pauldrach, Puls, Kudritzki 1986 – see lecture by Lamers in this volume). Once the basic hydrodynamic structure (essentially, the density and velocity as a function of radius) is determined, one solves in detail a NLTE radiative transfer, possibly together with the radiative equilibrium equation. This approach was pioneered by the Munich group (Gabler et al. 1989; Sellmaier et al. 1993), who also coined the term “unified models”. The name stresses a unification of a photosphere and wind; prior to this approach there were separate models for photospheres and for winds, so-called core-halo models.

Besides Munich models, there exists several other variants of unified model atmospheres. I do not present a review of these approaches (some topics are covered in other lectures (Lamers, this volume; Fullerton, this volume); I just briefly mention that various unified models are computed

- with or without self-consistent  $T(r)$ . That is, either the radiative equilibrium is solved exactly (e.g. Gabler et al. 1989; Hillier 1991); or the temperature structure is approximated for instance by the grey temperature structure (de Koter et al. 1993; Schaerer and Schmutz 1994);
- with or without Sobolev approximation in the wind;
- with or without metal line blanketing

## 5.2 Basic Equations of Classical Stellar Atmospheres

Let us summarize the basic equations of stellar atmospheres for the case of horizontally-homogeneous, plane-parallel, static atmosphere. This case is sometimes called the *classical stellar atmosphere problem*.

**Radiative transfer** equation. The most advantageous form of the transfer equation for the use in model atmosphere construction is either the usual first-order form, e.g. (29), which is then solved by the DFE method, or the second-order form with the variable Eddington factor,

$$\frac{d^2(f_\nu^K J_\nu)}{d\tau_\nu^2} = J_\nu - S_\nu . \quad (130)$$

It involves only the mean intensity of radiation,  $J_\nu$  (which is a function of only frequency and depth), but not the specific intensity (which is in addition a function of angle  $\mu$ ). In fact, it is the mean intensity of radiation which enters other structural equations, and therefore the mean intensities, not specific intensities, are to be taken as the atmospheric state parameters. An obvious numerical advantage is that instead of dealing with  $NF \times NA$  parameters describing the radiation field per depth ( $NF$  and  $NA$  being the number of discretized frequency and angle points, respectively) we have only  $NF$  parameters. A discretization of the depth variable, mentioned above, is done in such a way that depth points run from the “surface” depth, where  $\tau_\nu \ll 1$  for all frequency points, to a depth where  $\tau_\nu \gg 1$  for all frequencies (because the diffusion approximation, (109), is used for the lower boundary condition).

**Hydrostatic equilibrium** equation. This equation reads, recalling (7),

$$\frac{dP}{dz} = -\rho g , \quad (131)$$

where  $P$  is the total pressure. Introducing the Lagrangian mass  $m$ , defined as the mass in the column of a cross-section of  $1 \text{ cm}^2$  above a given point in the atmosphere,

$$dm = -\rho dz , \quad (132)$$

we obtain for the hydrostatic equilibrium equation simply

$$\frac{dP}{dm} = g , \quad (133)$$

which, since  $g$  is constant in a plane-parallel atmosphere, has a trivial solution,  $P(m) = mg + P(0)$ . In fact, this is the reason why one usually chooses  $m$  as the basic depth variable of the 1-D plane-parallel atmospheres problem. Nevertheless, it should be kept in mind that the total pressure is generally

composed of three parts, the gas pressure,  $P_{\text{gas}}$ , the radiation pressure,  $P_{\text{rad}}$ , and the turbulent pressure,  $P_{\text{turb}}$ , i.e.

$$P = P_{\text{gas}} + P_{\text{rad}} + P_{\text{turb}} = NkT + \frac{4\pi}{c} \int_0^\infty K_\nu d\nu + \frac{1}{2} \rho v_{\text{turb}}^2 , \quad (134)$$

where  $v_{\text{turb}}$  is the microturbulent velocity. The hydrostatic equilibrium equation may then be written as (neglecting the turbulent pressure)

$$\frac{dP_{\text{gas}}}{dm} = g - \frac{4\pi}{c} \int_0^\infty \frac{dK_\nu}{dm} = g - \frac{4\pi}{c} \int_0^\infty \frac{\chi_\nu}{\rho} H_\nu d\nu . \quad (135)$$

We may think of the r.h.s. of this equation as the *effective gravity acceleration*, since it expresses the action of the true gravity acceleration (acting downward, i.e. towards the center of the star) minus the radiative acceleration (acting outward). In other lectures (Lamers, this volume) we saw that this is the term which is crucial in the radiatively-driven wind theory.

**Radiative equilibrium** equation. This expresses the fact that the total radiation flux is conserved, see (8),

$$\int_0^\infty H_\nu d\nu = \text{const} = \frac{\sigma}{4\pi} T_{\text{eff}}^4 . \quad (136)$$

This equation may be rewritten, using the radiative transfer equation, as

$$\int_0^\infty (\kappa_\nu J_\nu - \eta_\nu) d\nu = \int_0^\infty \kappa_\nu (J_\nu - S_\nu) d\nu = 0 , \quad (137)$$

Notice that (137) contains the *thermal* absorption coefficient  $\kappa_\nu$ , not the total absorption coefficient  $\chi_\nu$ . This is because the scattering contributions cancel out. To illustrate this mathematically, let us take an example of electron scattering. The absorption coefficient for the process (see 141) is given by  $n_e \sigma_e$ ;  $\sigma_e$  being the electron scattering (Thomson) cross-section. The emission coefficient is then given by  $n_e \sigma_e J_\nu$ . As it is seen from (137), these two contributions cancel. This is also clear physically, because an absorption followed immediately by a re-emission of a photon does not change the energy balance of the medium, and therefore cannot contribute to the radiative equilibrium equation.

**Statistical equilibrium** equations; also sometimes called *rate equations*. These are in fact equations (6), where the collisional term is written explicitly,

$$n_i \sum_{j \neq i} (R_{ij} + C_{ij}) = \sum_{j \neq i} n_j (R_{ji} + C_{ji}) , \quad (138)$$

where  $R_{ij}$  and  $C_{ij}$  is the radiative and collisional rate, respectively, for the transition from level  $i$  to level  $j$ . The l.h.s. of (138) represents the total number

of transitions *out* of level  $i$ , while the r.h.s. represents the total number of transitions *into* level  $i$  from all other levels. The radiative rates are given by expressions analogous to those discussed for a two-level atom in Sect. 3.1 (notice that they depend on the radiation intensity), while the collisional rates are assumed to be given functions of temperature and electron density.

The set of rate equations for all levels of an atom would form a linearly dependent system. Therefore, one equation of the set has to be replaced by another equation. Usually, this is the *total number conservation* equation (or abundance definition equation),  $\sum_i n_i = N_{\text{atom}}$ , where the summation extends over all levels of all ions of a given species.

Two comments are in order. First, in practice there are only a limited number of levels of an atom/ion which are treated explicitly, i.e. for which the equation of the form (138) is actually written down and solved. These are usually low-lying levels. The remaining levels are typically treated in some approximate way, as, for instance, in LTE with respect to the ground state of the next ion (following Auer and Mihalas 1969), or with respect to the highest explicit level of the current ion. Another possibility is to express this contribution through the partition function (Hubeny 1988). In any case, the abundance definition equation has to be modified to read

$$\sum_{\text{explicit}} n_i + \sum_{\text{upper}} n_i = N_{\text{atom}} . \quad (139)$$

Second, the above abundance definition equation can replace the rate equation for *any* level. This level was usually taken, following Auer and Mihalas (1969), to be the ground state of the highest ion of the given species. However, a numerically more stable option is to choose a level which has the *highest population* of all the levels of the given species, as was suggested by Castor et al. (1992).

**Charge conservation** equation. This equation expresses the global electric neutrality of the medium,

$$\sum_i n_i Z_i - n_e = 0 , \quad (140)$$

where  $Z_i$  is the charge associated with level  $i$  (i.e. equal to 0 for levels of neutral atoms, 1 for levels for once ionized ions, etc.). The summation now extends over all levels of all ions of all species.

**Auxiliary definition equations.** There is a number of auxiliary expressions, like the definition equations of the absorption and emission coefficient,

$$\begin{aligned} \chi_\nu = & \sum_i \sum_{j>i} [n_i - (g_i/g_j)n_j] \sigma_{ij}(\nu) + \sum_i \left( n_i - n_i^* e^{-h\nu/kT} \right) \sigma_{i\kappa}(\nu) \\ & + \sum_\kappa n_e n_\kappa \sigma_{\kappa\kappa}(\nu, T) \left( 1 - e^{-h\nu/kT} \right) + n_e \sigma_e , \end{aligned} \quad (141)$$

where the four terms represent, respectively, the contributions of bound-bound transitions (i.e. spectral lines), bound-free transitions (continua), free-free absorption (also called the inverse brehmstrahlung), and of electron scattering. Other scattering terms, like for instance the Rayleigh scattering, may also be added if appropriate to the problem at hand. Here,  $\sigma(\nu)$  are the corresponding cross-sections; subscript  $\kappa$  denotes the “continuum”, and  $n_\kappa$  the ion number density. The negative contributions in the first three terms represent the stimulated emission (remember, stimulated emission is treated as negative absorption). There is no stimulated emission correction for the scattering term, since this contribution exactly cancels with ordinary absorption (for an illuminating discussion, see Shu 1991). Finally, notice that the relation between the bound-bound cross section  $\sigma_{ij}(\nu)$  and previously introduced quantities (the Einstein coefficients and the absorption profile) is simply  $\sigma_{ij}(\nu) = (h\nu_0/4\pi)B_{ij}\phi(\nu)$ .

Analogously, the thermal emission coefficient is given by

$$\eta_\nu = (2h\nu^3/c^2) \left[ \sum_i \sum_{j>i} n_j (g_i/g_j) \sigma_{ij}(\nu) + \sum_i n_i^* \sigma_{i\kappa}(\nu) e^{-h\nu/kT} + \sum_\kappa n_e n_\kappa \sigma_{\kappa\kappa}(\nu, T) e^{-h\nu/kT} \right]. \quad (142)$$

The three terms again describe the bound-bound, bound-free, and free-free emission processes, respectively.

These equations should be complemented by expressions for the relevant cross-sections, definition of LTE populations, and other necessary expressions. The resulting set forms a highly-coupled, highly non-linear system of equations. The equations and corresponding quantities that are determined by them are summarized in the Table 1.

**Table 1.** Summary of classical stellar atmosphere equations and state parameters

Equation	Corresponding state parameter
Radiative transfer	Mean intensities, $J_\nu$
Radiative equilibrium	Temperature, $T$
Hydrostatic equilibrium	Total particle density, $N$
Statistical equilibrium	Populations, $n_i$
Charge conservation	Electron density, $n_e$

### 5.3 LTE-grey Model: A Tool to Understand the Temperature Structure

Before discussing the methods and results of solving the full stellar atmosphere problem, it is very instructive to consider an extremely simplified case of the so-called LTE-grey model. Although these models have not been used to describe a real stellar atmosphere for more than four decades, they are still very useful because i) they provide a beautiful pedagogical tool to understand an interplay between radiative equilibrium and radiative transfer, thus to understand a behavior of temperature as a function of depth; and ii) they provide an excellent starting solution for iterative methods to construct more sophisticated models.

The basic assumption of these models is that the absorption coefficient is independent of frequency,

$$\chi_\nu \equiv \chi \quad . \quad (143)$$

In reality, one uses some frequency-averaged opacity, usually the Rosseland mean opacity, (71). The other basic assumption is that of LTE,  $S_\nu = B_\nu$ . The radiative equilibrium equation thus reduces to

$$J = B \quad , \quad (144)$$

where the quantities without the frequency subscript  $\nu$  are understood as frequency-integrated quantities,

$$J = \int_0^\infty J_\nu d\nu \quad ; \quad B = \int_0^\infty B_\nu d\nu = \sigma T^4 \quad . \quad (145)$$

The second equation of the problem, the radiative transfer equation (actually, its second moment), reads

$$\frac{dK}{d\tau} = H \quad \implies \quad K(\tau) = H \cdot \tau + \text{const} \quad , \quad (146)$$

because  $H$  is constant with depth, as follows from the radiative equilibrium. The constant in the above equation is equal to  $K(0)$ . Invoking, for simplicity, the Eddington approximation –  $K = J/3$ , and  $K(0) = (2/3)H$  – we obtain (recall that the flux is given by  $F = 4H$ ),

$$J(\tau) = \frac{3}{4}F \cdot \left( \tau + \frac{2}{3} \right) \quad , \quad (147)$$

We know that the total flux,  $F$ , is specified through the effective temperature,  $F = \sigma T_{\text{eff}}^4$ . Combining (145) and (147) together, we obtain

$$T^4 = \frac{3}{4} T_{\text{eff}}^4 \left( \tau + \frac{2}{3} \right) \quad , \quad (148)$$

There exists an elegant analytic solution of the general grey atmosphere problem which yields an analogous expression for the temperature as (148), only

the constant  $2/3$  is replaced by a function  $q(\tau)$ , called *Hopf function*. It is a smoothly varying function of optical depth, with  $q(0) = 0.577$ ;  $q(\infty) = 0.71$ , which is not very far from the Eddington approximation value.

An important point to realize is that the grey temperature structure follows just from the radiative transfer equation and the radiative equilibrium equation. The hydrostatic equilibrium equation does not enter this derivation. In other words, the temperature in a grey atmosphere, as a function of mean optical depth, does not depend on the surface gravity. However, the hydrostatic equation determines the relation between the averaged optical depth and the geometrical coordinate ( $m$  or  $z$ ).

We see that the temperature is a monotonically increasing function of optical depth. Why this is so? It is easy to understand it in physical terms. The condition of radiative equilibrium stipulates that the *total* radiation flux is constant with depth in the atmosphere. However, the radiation flux measures the *anisotropy* of the radiation field (i.e. the flux would be zero for perfectly isotropic radiation). We know from the transfer equation, and in particular from the diffusion approximation, that the anisotropy decreases with increasing depth in the atmosphere. The only way how to maintain the constant flux in spite of decreasing anisotropy of radiation is to increase the total energy density of radiation (proportional to  $J$ ), i.e. the temperature (recall that  $J = S = B = \sigma T^4$ ).

The fact the integrated  $J$  is equal to integrated  $B$  at all depths  $\tau$  does not necessarily mean that the frequency-dependent  $J_\nu$  has to be equal to  $B_\nu$  for all frequencies. In fact, we should expect that there should be a frequency range for which  $J_\nu > B_\nu$ , i.e.  $J_\nu - B_\nu > 0$ ; these regions may be called “heating” regions; while at the rest of frequencies  $J_\nu < B_\nu$ , i.e.  $J_\nu - B_\nu < 0$ ; these regions may be called “cooling” regions. Remember,  $J$  is proportional to the number of photons absorbed per unit optical depth, while  $S = B$  to the number of photons emitted per unit optical depth. Thus, for instance,  $J_\nu > B_\nu$  means that more photons are absorbed than emitted at frequency  $\nu$ ; the energy of extra absorbed photons must then increase the internal energy, i.e. the temperature, of the medium.

Which frequency regions are the heating ones, and which are the cooling ones? In the case of an LTE-grey atmosphere, the answer is easy. Let us first write down some useful expressions. From the general expression for the Plack function, (22), we may easily derive two limiting expressions: In the high frequency limit,  $(h\nu/kT) \gg 1$ , we obtain the *Wien* form,

$$B(\nu, T) \approx \frac{2h\nu^3}{c^2} \exp(-h\nu/kT) , \quad (149)$$

while the low-frequency limit,  $(h\nu/kT) \ll 1$ , is called the *Rayleigh-Jeans* tail,

$$B(\nu, T) \approx \frac{2k\nu^2}{c^2} T . \quad (150)$$

Another important expression is the Eddington-Barbier relation for the mean intensity at the surface, which may be derived from the Eddington-Barbier relation (41) integrated over angles,

$$J(0) = \frac{1}{2}S(\tau=1) \quad (151)$$

Let us consider the surface layer of a grey atmosphere. If the frequency  $\nu$  is “large”, i.e. in the Wien regime, then a decrease of the local temperature between  $\tau = 1$  and the surface ( $\tau = 0$ ), translates into a large decrease of  $B_\nu(T(\tau))$ , because for large frequencies the Planck function is very sensitive to  $T$  – see (149). In other words,  $B$  at the surface may be significantly (even orders of magnitude) lower than  $B$  at  $\tau = 1$ . Since the mean intensity at the surface is about one half of  $B(\tau=1)$ , it is clear that  $J_\nu(0) > B_\nu(0)$  for these frequencies. The large frequencies are therefore the “heating” frequencies.

In contrast, for low frequencies (the Rayleigh-Jeans tail),  $B$  is linearly proportional to  $T$ . We know from the  $T(\tau)$  relation for a grey atmosphere that  $T(0) \approx 0.8 T(\tau=1)$ . The factor  $1/2$  from the Eddington-Barbier relation will now dominate, so we get  $J_\nu(0) = (1/2)B_\nu(\tau=1) < B_\nu(0)$ . Consequently, the low frequencies are the “cooling” frequencies. One can make these considerations more quantitative, but this is not necessary; the only important point to remember is that the high-frequency part of the spectrum is responsible for heating, while the low-frequency part is responsible for cooling.

**Two-step grey model.** The above considerations are interesting, but not particularly useful for a purely grey atmosphere. They are, however, very helpful if we consider an atmosphere with some simple departures from the greyness. Let us consider a *two-step grey* model, i.e. with the opacity given as a step function,  $\chi_\nu = \chi$  (the original grey opacity) for  $\nu < \nu_0$ , and  $\chi_\nu = a\chi$  for  $\nu \geq \nu_0$ , with  $a \gg 1$ , i.e. with a large opacity for high frequencies (one may visualize this as a schematic representation of a strong continuum jump, for instance the Lyman discontinuity). We will denote the original optical depth as  $\tau^{\text{old}}$ , and the new one (for  $\nu \geq \nu_0$ ), as  $\tau_\nu^{\text{new}}$ . Let us further assume that the frequency  $\nu_0$  is high enough to be in the range of “heating” frequencies.

What are the changes of the temperature structure with respect to the original grey temperature distribution implied by the opacity jump? We will consider separately the surface layers  $\tau \approx 0$ , and the deep layers.

*The surface layers.* Since the opacity for  $\nu \geq \nu_0$  is much larger than the original opacity, we may neglect the contribution of the latter to the radiative equilibrium integral, so the modified radiative equilibrium equation becomes

$$\int_{\nu_0}^{\infty} J_\nu d\nu = \int_{\nu_0}^{\infty} B_\nu d\nu \quad , \quad (152)$$

which, together with the Eddington-Barbier relation  $J_\nu(0) = B_\nu(\tau_\nu^{\text{new}} = 1)/2$  yields for the new surface temperature,  $T_0$ , the expression

$$(1/2) \int_{\nu_0}^{\infty} B_\nu (T(\tau_\nu^{\text{new}} = 1)) d\nu = \int_{\nu_0}^{\infty} B_\nu(T_0) d\nu \quad , \quad (153)$$

from which follows that  $T_0 < T(\tau_\nu^{\text{new}} = 1)$ . Since the temperature at  $\tau_\nu^{\text{new}} = 1$  must be close to the original temperature at the surface (recall that  $\tau_\nu^{\text{new}} \gg \tau^{\text{old}}$ ), the new temperature at the surface is *lower* than the original surface temperature, which gives rise to the term *surface cooling* effect.

The above derivation was more or less a mathematical one. But, in physical terms, why do we get a cooling? This is simply because by *adding opacity in the heating portion* of the spectrum, we effectively suppress this heating. Therefore, we obtain a cooling. These considerations also suggest that by adding an additional opacity in the cooling, i.e. the low-frequency, part of the spectrum, we may actually get a *surface heating* of the atmosphere.

*The deep layers.* It is intuitively clear that the atmospheric layers which are optically thick in all frequencies will be little influenced by the additional opacity jump. However, an interesting region is the one which is opaque for large frequencies ( $\nu \geq \nu_0$ ), (i.e.  $\tau_\nu^{\text{new}} \gg 1$  for these frequencies), while still transparent for the original opacity,  $\tau^{\text{old}} < 1$ . Since the optical depth is large for  $\nu \geq \nu_0$ ,  $J_\nu \approx B_\nu$  for these frequencies, and therefore the monochromatic flux is close to zero. The condition of radiative equilibrium at those depths may be written as  $J' = B'$ , where the primed quantities are defined as partial integrals, e.g.  $J' = \int_0^{\nu_0} J_\nu d\nu$ , and analogously for  $B$ . From the radiative transfer equation and the Eddington approximation, we have  $dJ'/d\tau = 3H$  (not  $H'$ ; or, better speaking,  $H' = H$ , because there is no flux for  $\nu \geq \nu_0$ ). We may formally write  $J' = \sigma' T^4$ , and by repeating the same procedure as in deriving the original grey temperature structure, we obtain

$$T^4 = (3/4)(\sigma/\sigma') T_{\text{eff}}^4 (\tau + 2/3) . \quad (154)$$

We have  $\sigma' < \sigma$ , because  $J' < J$ . This is simply because the energy density of radiation for  $\nu < \nu_0$  is smaller than the total energy density. Therefore, the new temperature is larger than the original one. Consequently, the phenomenon is called the *backwarming effect*.

Again, what is the explanation of this effect in physical terms? By adding opacity, the flux in the high-opacity part *drops*. Therefore, the flux in the rest must *increase* in order to keep the total flux constant. However, the only way how to accomplish it in LTE is to increase the temperature gradient, and therefore the temperature itself in the previously flat  $T(\tau)$  region.

One may wonder why we spend so much time with an admittedly crude and unrealistic model, such as a simple two-step grey model. However, it should be realized that the above discussed phenomena of surface cooling and backwarming are quite general, and are not at all limited to a grey approximation. In any model, including sophisticated NLTE models (see Sect. 5.5), there are always frequencies which cause heating and those which cause cooling. Any process which changes opacity/emissivity in those regions changes the overall balance and therefore influences the temperature structure. In the NLTE models, there are typically several intervening or competing mechanisms, but the fundamental physics behind the temperature structure is basically the same as in the case of the grey model. Likewise, the mechanism of

backwarming is quite general. The beauty of the grey model is that one may describe all these phenomena by a simple analytical model.

#### 5.4 LTE and NLTE Model Atmospheres

**LTE models.** Constructing LTE model stellar atmospheres is now a more or less standard procedure. It consists in solving simultaneously basic structural equations (130) - (142), where (138) is replaced by the Saha-Boltzmann distribution, (11) and (12). Consequently, the absorption and emission coefficients are known functions of temperature and electron density, i.e. they are given *locally*. Nevertheless, there is still a non-local coupling of radiation field and material properties via the radiative equilibrium equation (and, to a smaller extent also the hydrostatic equilibrium equation, via the radiation pressure term), which has to be dealt with.

I will not discuss this topic here in any detail. I just mention that the field of LTE model atmospheres is completely dominated by the Kurucz model grid (Kurucz 1979; 1994), and by his computer program ATLAS (Kurucz 1970; 1994). Yet, there are several independent computer programs, designed specifically for very cool stars – Gustafsson et al. (1975); Tsuji (1976); Johnson et al. (1977); Allard and Hauschildt (1995); to name just few.

**NLTE models.** Why do we expect that departures from LTE may be important in stellar atmospheres? As explained above, departures from LTE arise when the radiative rates dominate over the collisional rates. These conditions typically occur at high temperatures and low densities. The higher the effective temperature, and the stronger the radiation field, the deeper in the photosphere we may expect departures from LTE. We also anticipate that the departures will be largest at frequencies with highest opacities (EUV, cores of strong lines). When the opacity is large, the observed spectrum will be formed higher in the atmosphere where the density of the material is low. Therefore, NLTE models are most important for interpreting observed spectra of hot stars (O, B, A stars, typically  $T_{\text{eff}} > 10\,000$  K) and of supergiants, i.e. the intrinsically brightest stars.

However, the most important point to realize is that *for a star of any spectral type, there is always a wavelength range, and correspondingly a layer in the atmosphere, where NLTE effects are important*. Yet, the meaning of the assessment “NLTE effects are important” is somewhat arbitrary. The point is that a precise definition of this term should in principle involve the desired accuracy of the predicted spectrum. For instance, if one requires an accuracy of, say, 10% in the predicted flux in the optical and UV spectrum for a main-sequence B star, then one may say that LTE models are sufficient. However, NLTE models would be necessary if one requires an accuracy of, say 2-5 %; and NLTE models would still be necessary if one requires an accuracy of 10-20% for the same star in the EUV spectrum range (wavelengths below the Lyman limit, i.e. 912 Å).

How one calculates NLTE model atmospheres? The realization that the nonlocal coupling of physical quantities is extremely important led Auer and Mihalas (1969) to develop the complete linearization (CL) method to solve the set of discretized structural equations. This is a very robust method based on the Newton-Raphson scheme. All equations are linearized and are treated on the same footing, allowing a fully consistent treatment of all couplings. The method has brought an enormous progress in the modeling, and in fact has opened a new era in the stellar atmospheres theory. However, a high price had to be paid. Because of the need to invert individual block matrices of the general Jacobi matrix of the system, the computer time increases with the cube of the number of unknowns. Therefore, it was possible to consider only a limited number of atomic levels and opacity sources (lines). Typically, about 10 to 15 energy levels were allowed to depart from their LTE populations; only a few lines were included explicitly, and the radiative transfer was solved at typically 100 frequencies. It also became clear very soon that dealing with millions of lines within this framework would be out of the question regardless of how rapidly the computer technology might progress.

However, already the early simplified NLTE models have demonstrated that departures from LTE form an essential feature in interpreting the spectra of hot stars (for a review, see Mihalas 1978, and Kudritzki and Hummer 1990). In the same period, the importance of metal line blanketing on the atmospheric structure was demonstrated numerically, and a widely used grid of LTE line-blanketed model atmospheres was constructed (Kurucz 1979). Since then, a debate ensued as to what kind of model atmospheres is more adequate: metal line-blanketed LTE models or NLTE models without blanketing? Models accounting for both metal line-blanketing effects and departures from LTE were then deemed an unreachable dream.

The dream had slowly started to come through in the 1980's with the advent of the ALI method. The first who applied the ALI idea to the model stellar atmosphere construction was Werner (1986; 1987; 1989). He has shown that the method has a great potential, because it effectively eliminates the radiation intensities from the set of model unknowns. One is therefore able to consider many more frequency points, and consequently many more atomic transitions, in the model construction. Moreover, the method was found to be more stable than the complete linearization method in many cases.

A disadvantage of the ALI scheme is that it sometimes converges rather slowly. This is easy to understand, since the information about changes in state parameters is lagged, i.e. is communicated to the rest of the state parameters only in the subsequent iteration. Nevertheless, the time per iteration is very small. Moreover, the speed of convergence can be accelerated by predicting better estimates of the solution using the acceleration of convergence techniques (see Sect. 4.3). In contrast, the CL method requires only a small number of iterations, because it is a global method with an almost quadratic convergence. The time per iteration may however be enormous.

What was therefore needed was a method which would combine the advantages of both these methods; namely the convergence rate (i.e. the number of iterations required to reach a given accuracy) being virtually as high as for the standard CL method, while the computer time per iteration is almost as low as for the standard ALI method. This was exactly what was achieved by developing the so-called hybrid complete linearization/accelerated lambda iteration (CL/ALI) method (Hubeny and Lanz 1995). The method formally resembles the standard complete linearization; the only difference being that the radiation intensity at selected frequency points is not explicitly linearized; instead, it is treated by means of the ALI approach.

### 5.5 Line Blanketing

The term line blanketing describes an influence of thousands to millions of spectral lines on the atmospheric structure and predicted emergent spectrum. The line blanketing influences not only the emergent spectrum (the so-called line blocking), but also, and more importantly, the atmospheric structure (the backwarming and the surface cooling effects). Although the ALI-based methods have opened the way to attack this problem, the enormous complexity of the iron-peak elements (i.e., we have to account for hundreds of energy levels and millions of line transitions per ion) still precludes using direct methods which were successfully used for light elements (He, C, N, O, etc.).

Statistical methods are therefore necessary. The idea is to avoid dealing with all individual energy levels of complicated metal species. Instead, several states with close enough energies are grouped together to form a so-called “superlevel”. The basic assumption is that all individual levels within the same superlevel share the same NLTE departure coefficient; in other words, the individual levels forming a superlevel are in Boltzmann equilibrium with each other. This idea was pioneered by Anderson (1989). The transitions between individual superlevels, called “superlines”, are treated by means of two possible approaches:

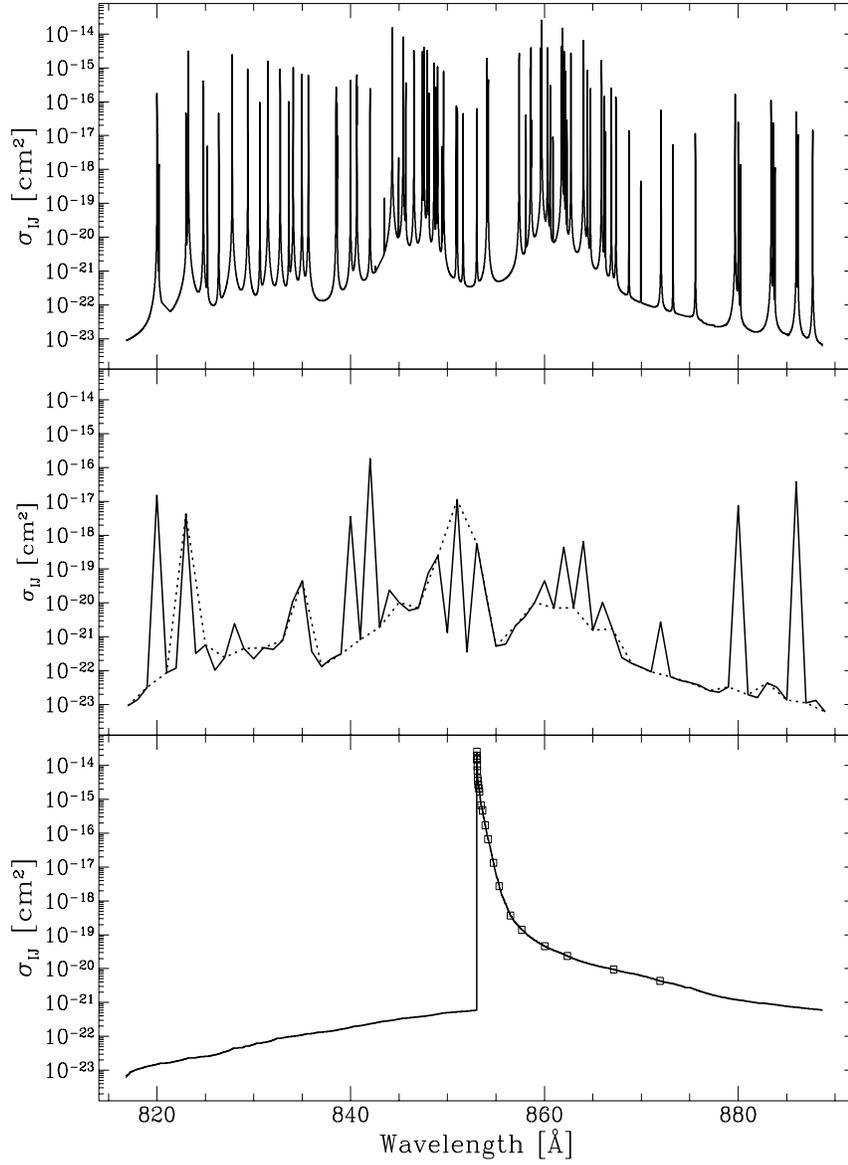
- *Opacity Distribution Functions* (ODF). The idea is to resample a complicated frequency dependence of the superline cross-section to form a monotonic function of frequency; this function is then represented by a small number of frequency quadrature points (Anderson 1989; Hubeny and Lanz 1995).
- *Opacity Sampling* (OS). The idea is a simple Monte Carlo-like sampling of frequency points of the superlevel cross-section (Anderson 1991; Dreizler and Werner 1993). The advantage of this approach is that it can easily treat line blends and overlaps; the disadvantage is that one has to be very careful to choose a sufficiently large number of frequency points, since otherwise the representation may be inaccurate. Indeed, the line cores, which represent the region of maximum opacity, are relatively narrow. Considering too few frequency points may easily lead to missing many important line cores.

These two approaches are illustrated in Fig. 5. We consider the superline between the superlevels 1 and 13 of Hubeny and Lanz (1995) model of Fe III. The detailed cross-section (upper panel) has been computed for some 16 000 internal frequency points. The dotted line in the middle panel represents the Opacity Sampling by 37 (equidistant) wavelength points, while the number of points is doubled for the full line. This shows that unless a large number of frequency points is considered, the OS representation may be quite inaccurate since practically all strong lines are missed. Finally, the lower panel shows the Opacity Distribution Function representation. With 24 points only, we have already a fairly accurate representation of the resampled cross-section to be used in model atmosphere construction.

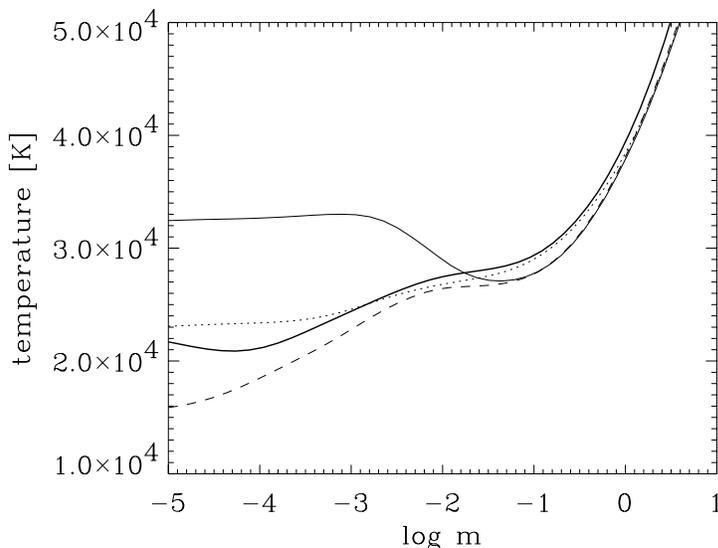
The first NLTE model atmospheres including iron-line blanketing were presented by Anderson (1989), Anderson and Grigsby (1991), Dreizler and Werner (1993), and Hubeny and Lanz (1995). The strategy for computing line-blanketed model atmospheres is as follows. Hydrogen, helium, and the most important light metals (C, N, O, possibly others) are represented by detailed atomic models, and all the individual lines are treated separately. This involves of the order of 100 atomic levels, and up to 1 000 lines, which are represented by several thousands of frequency points. The heavy (iron-peak) metals are treated by means of the statistical, ODF or OS, approach. Since the dominant opacity is provided by iron and nickel, we either neglect all the other iron-peak elements, or group all of them together to form an averaged iron-peak element (as suggested first by Anderson 1989).

Finally, I will show how the line blanketing influences the temperature structure of an atmosphere. As an example, let us take a model with  $T_{\text{eff}} = 35\,000$  K, and  $\log g = 4$ , which corresponds to a main-sequence O-star. Figure 6 shows the temperature as a function of depth (expressed as column mass in  $\text{g cm}^{-2}$ ). We consider several NLTE model atmospheres, a) H-He LTE model, b) H-He NLTE model, c) NLTE model with light elements only (H-He-C-N-O-Si); and d) fully blanketed NLTE model (H-He-C-N-O-Si-Fe-Ni). All models consider all lines originating between explicit levels of all species that are taken into account.

The behavior of temperature is easily explained by a reasoning analogous to that put forward in the preceding subsection. The frequency region above the Balmer limit (i.e.  $\lambda < 3648$  Å) is the “heating region”. Therefore, adding an opacity there causes a surface cooling. This explains the cooling in LTE H-He model (caused mainly by the hydrogen and He II Lyman and Balmer lines), and also the additional cooling in the H-He-C-N-O-Si model (which is caused mainly by the C IV resonance doublet at  $\lambda 1550$  Å). Similarly, the additional opacity in the heating region causes the heating of deeper layers, the so-called backwarming effect. Indeed, it is clearly seen that while the lines of light elements cause only a modest backwarming (in the layers at  $\log m \approx -1$  and deeper), the Fe and Ni lines, being quite numerous, cause an appreciable heating in these layers.



**Fig. 5.** An illustration of various numerical treatments of a typical superline. Upper panel: the detailed cross-section; Middle panel: the Opacity Sampling representation; Lower panel: the Opacity Distribution Function representation. Small squares indicate the points used to represent this ODF in model atmosphere calculations



**Fig. 6.** Temperature structure for four model atmospheres with the same parameters:  $T_{\text{eff}} = 35\,000$  K,  $\log g = 4$ . Thick line: fully blanketed NLTE H-He-C-N-O-Si-Fe-Ni model; dashed line: NLTE model with light elements (H-He-C-N-O-Si); thin line: NLTE H-He model; dotted line: LTE H-He model.

Now, how it is possible that few lines (of H, He, or light elements) are able to cause a significant surface cooling, while a large number of lines is needed to get a significant backwarming? Again, this is explained by employing the two-step grey model considerations. Let us take equation (152). It shows that a strong opacity source completely dominates the radiative equilibrium integral, so that the other frequency regions become unimportant. The original two-step grey model considers the strong opacity source to extend from  $\nu_0$  to infinity; however, the essence remains the same if the strong opacity source is just one line, or few strong lines. In the case of one dominant line, the radiative equilibrium integral reduces to, in analogy to (152),

$$\bar{J} = S^{\text{L}} , \quad (155)$$

which follows from (137), (72), (73), and (78). In LTE, we get surface cooling due to the exactly same reasons as in the two-step grey model (LTE forces  $S$  to be equal to  $B$ , and  $B$  is forced to be equal to  $\bar{J}$  at the surface, which is low). In NLTE, the cooling effect may be even stronger, because the two-level source function, (82), implies that  $\bar{J} = S^{\text{L}} = B$ , and we know from Sect. 3.1.

that the two-level atom source function decreases significantly towards the surface.

On the other hand, the backwarming effect is primarily caused by blocking of the flux by additional opacity sources; the more extended (in the frequency space) the blocking is, the larger the backwarming (recall (154) and subsequent discussion). The actual strength of a line does not matter so much as soon as it is able to efficiently block the flux. Therefore, a single narrow and very strong line is quite efficient in the surface cooling, but relatively inefficient in the backwarming.

There is still one remaining puzzling feature: Why, in view of all what was said above, we obtain a large *surface heating* in the NLTE H-He models? This was indeed a big surprise when the effect was first discovered by Auer and Mihalas (1969), who have also provided its physical explanation. The effect is related, but not equivalent, to another, previously discovered NLTE surface temperature rise, called the Cayrel mechanism (Cayrel 1963).

The explanation of the Auer-Mihalas temperature rise goes as follows: It is true that lines always cause a surface cooling. However, in NLTE, a line radiation also influences the atomic level populations. From Sect. 3.1 we know that the main effect of line transfer is to overpopulate the lower level of a line transition. Considering Lyman and Balmer lines thus gives rise to an overpopulation of the hydrogen  $n = 1$  and  $n = 2$  states, and consequently to increasing the efficiency of the Lyman and Balmer continua. Since they are heating continua, this leads to an additional heating at the surface. There is a competition between this heating and traditional surface cooling caused by the Lyman and Balmer lines, but in the present case the indirect heating wins.

However, interesting as it may be from the theoretical point of view, the indirect heating due to the hydrogen (and to a lesser extent the He II) lines is in reality usually wiped out by the effect of metal lines (as it is in the case displayed in Fig. 6). Nevertheless, the Auer-Mihalas heating survives for metal-poor atmospheres, where it may give rise to observable effects in the hydrogen line profiles (e.g. for hot DA white dwarfs – see Lanz and Hubeny 1995 for a discussion and original references).

Finally, I stress that the behavior of temperature at the surface should not be overinterpreted. It only influences observed spectrum features which correspond to the strongest opacity sources, like the very cores of strongest lines (e.g. the C IV resonance lines in the present case). Yet, these features may in reality be more influenced by a stellar wind, which is neglected in the hydrostatic models anyway. Therefore, the most important effect of line blanketing is its influence on temperature in the deeper layers.

## 6 Using Model Atmospheres to Analyse Observed Spectra

So far, we have been mostly concerned with the question of how the model stellar atmospheres are constructed. In this chapter, we will discuss another, and equally important, question of how these models are used to address general astrophysical questions.

### 6.1 A Scheme of Spectroscopic Diagnostics

As was stated before, the observed spectrum is practically the only information about a star we have. The process of deducing stellar properties from its spectrum is therefore called *spectroscopic diagnostics*. This is a multi-step process with many interlinked steps. It is schematically displayed in Fig. 7.

The basic step is `INPUT PHYSICS`. By this term we mean a selection of the basic physical assumptions under which the medium is being described (i.e., which processes and phenomena are neglected; which equilibrium conditions are assumed to hold, etc.). The basic assumptions then determine the equations to be solved. They also tell us what are the basic input parameters of the model construction. For instance, when adopting the assumption of a plane-parallel atmosphere in the hydrostatic and radiative equilibrium, the input model parameters are the effective temperature, surface gravity, and chemical composition. These parameters are “basic” from the point of constructing model atmospheres, yet they are related to other parameters which may be viewed as more fundamental, like stellar mass, radius, and luminosity. The latter parameters are fundamental if one considers a certain instant of the stellar life. Yet, taking into account more extended input physics (i.e., adding the stellar evolution theory), we may then consider even more fundamental parameters like the initial stellar mass, initial composition, and the age.

Likewise, going to more complex models, like for instance the unified photosphere–wind models, we have different input parameters depending on the level of physical description we adopt. In a simple theory we have, in addition to the stellar mass, radius, and luminosity, two additional input parameters – the mass loss rate and the wind terminal velocity (see Lamers, this volume). Yet, in a more involved physical picture we may come up with a relation between the wind parameters and other parameters.

Sometimes the additional input parameters make up for the lack of adequate physics. Typical examples are the so-called microturbulent velocity which is often used for describing short-scale non-thermal motions; or the mixing length parameter used in the mixing-length theory of convection. An example from a somewhat different yet related field is the  $\alpha$ -parameter for describing a turbulent viscous dissipation in accretion disks. All such input parameters are convenient parameters which allow us to construct models even if we do not really know what is going on. Their values are constrained

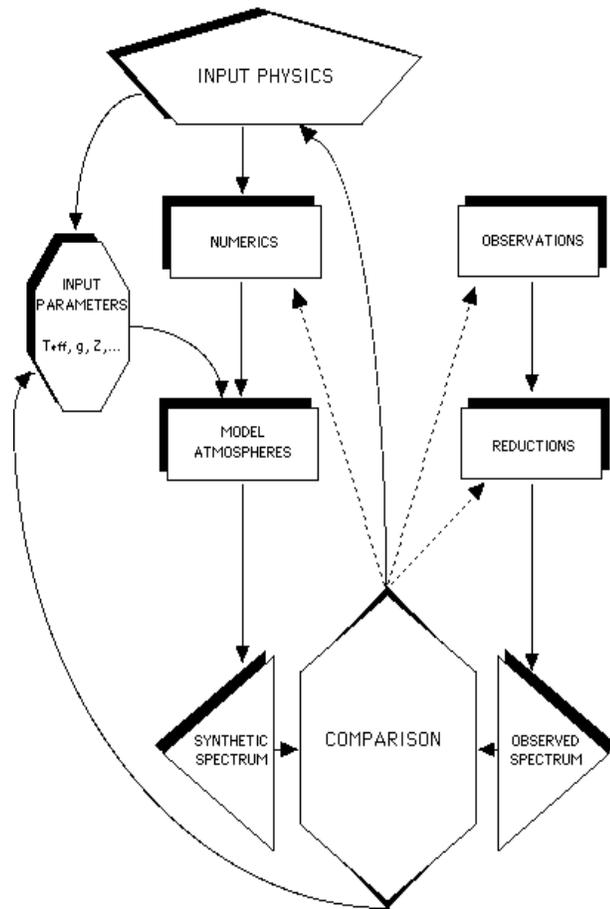


Fig. 7. A sketch of the individual steps of astrophysical spectroscopic diagnostics and their interconnections

by agreement between theoretical predictions and observations; it is generally believed that a more fundamental physics will either determine their values from other structural parameters, or will get rid of them completely. Nevertheless, there are cases where such parameters are very useful, for instance when a “better physics” would require an enormously complicated and time-consuming modeling.

The input physics, which tells us which equations are to be solved, thus influences profoundly the next step, NUMERICS. By this term we understand all the work necessary to develop a code for computing model stellar atmospheres. This involves adopting appropriate methods provided by numerical mathematics or, often, developing new methods suited to a particular modeling purpose (a good example being the ALI method). This also involves a lot of computer programming and, the most time-consuming part, code debugging and testing.

Having developed a stellar atmosphere code, one may proceed to the next step, MODEL ATMOSPHERES. It is depicted in Fig. 7 as a distinct step from numerics, despite the fact that it could have been a part of the Numerics box. Usually, a stellar atmosphere code contains a large number of various numerical options and tricks. One usually needs a lot of experience to cope successfully with various numerical problems (typically a slow convergence or divergence of iterations), and to find proper options to coax the code to work. Sometimes the author of the code builds a grid of models him or herself (typical example being Bob Kurucz), but it is still useful that a code itself is being available to the whole community. This is because the number of internal input parameters may be enormous to make it reasonable to construct a sufficiently dense, all-purpose grid. Many codes for stellar atmospheric modeling are indeed publicly available.<sup>1</sup>

The last step of the “theoretical” branch of the spectroscopic diagnostic procedure is spectrum synthesis, which yields the main product, the SYNTHETIC SPECTRUM. It will be discussed in detail in the next section.

It should be realized that not all of the above steps have to be done in analyzing a particular object. One may work, for instance, with a pre-calculated grid of model atmospheres and construct only synthetic spectra. One may even work with an existing grid of synthetic spectra; one then avoids the theoretical part completely.

I will not discuss the other, “observational”, branch of diagnostics. The main steps are taking the rough data (by ground-based or space-based instruments) – the step OBSERVATIONS – and subsequent DATA REDUCTIONS. The final product is a well-calibrated OBSERVED SPECTRUM.

Now comes the crucial part of the spectroscopic diagnostics, the COM-

---

<sup>1</sup> The most extended collection of existing modeling codes is maintained on the CCP7 (Collaborative Computer Project No. 7 on the Analysis of Astronomical Spectra) library – Jeffery (1992). The library is also available via WWW on the address <http://star.arm.ac.uk/ccp7/>

PARISON of the observed and theoretical spectra. The actual procedure to be performed in this step is discussed in detail in Sect. 6.3. Here, I will only stress the significance of this step in the global context. Above all, the comparison determines the values of basic input parameters. One may proceed iteratively: after an initial guess of input parameters and constructing the first synthetic spectrum, the comparison step suggests new values of basic input parameters, which are then used for constructing a new model atmosphere and new synthetic spectrum, and the process is repeated. Alternatively, one may first construct a grid of synthetic spectra around the most probable values of basic parameters, and to determine their final values by some sort of  $\chi^2$ -fitting.

However, the most important point to realize is that the comparison step does not merely serve to determine values of model input parameters. It may happen that we are not able to match observations for any combination of input parameters. Then we have to go back to the input physics step, and revise the basic physical assumption under which the models were constructed. This may of course lead to a revision (or even to a rewriting) of the computer program and consequently to reconstructing the model grid. But by this connection we actually learn the most important part of all – the physics.

In Fig. 7, the dashed line drawn from the comparisons step to the numerics step is meant to indicate that lack of agreement between observations and theory does not have to be caused by an inadequacy of adopted physical description, but also by an inadequate numerical treatment of otherwise correct equations. The most trivial example of this phenomenon is a hidden bug in the program, which does not influence the results in some cases (usually those used for testing), but may have a dramatic effect in others. Also, this may indicate an insufficient accuracy or inadequacy of adopted numerical method(s) for solving a given set of equations, for inverting matrices, etc. Another example are various numerical approximations for atomic parameters. (e.g. polynomial expansions for transition cross-section or collisional rates, which are derived for a certain parameter range, but are not applicable for others, etc., etc.)

Finally, an unsuccessful match of observations and theory may reveal that the observations are at fault. For instance, an UV spectrum of a star may be contaminated by a contribution from a close object not seen in the optical region. Also, the data reduction step may be at fault – for instance a persistent mismatch of observations and models for various objects may lead to a discovery of an error in flatfield corrections, or an unexpected presence of scattered light in the spectrograph, etc.

## 6.2 Spectrum Synthesis

Ideally, there should be a one-to-one correspondence between a model atmosphere (the structure), and the synthetic spectrum. In other words, to every model atmosphere there should correspond a unique emergent spectrum. However, in the real life the model structure is always computed using

a simplified treatment of chemical species – some of them do not influence the atmospheric structure significantly and are therefore omitted (like, for instance, Li, etc.), or are treated in a simplified way when constructing the model.

However, one needs to include all the opacity sources available, lines and continua, when producing the synthetic spectrum. One can easily afford that. Since the temperature, electron density, and atomic level populations are specified by the model, the calculation of the synthetic spectrum consists of a simple wavelength-by-wavelength formal solution of the radiative transfer equation, with the absorption and emission coefficients given by (141) and (142). However, the summation over line transitions may actually include hundreds or even thousands of individual spectral lines contributing at a single wavelength point. This feature, and the very fact that we have to deal with literally millions of lines, make an efficient coding of this problem non-trivial. The most widely used general purpose codes of this kind are SYNTHE (Kurucz 1994), which is designed to produce spectra for Kurucz model atmospheres; and SYNSPEC (Hubeny, Lanz, and Jeffery 1994), which calculates spectra for NLTE models created by TLUSTY, but works for Kurucz models as well.

Another important point: an input model atmosphere is constructed assuming certain abundances of chemical species. In order to be strictly consistent, one would have to consider the same abundances in the spectrum synthesis as well. However, due to the same reasons as put forward above, it is permissible to use different abundances in the spectrum synthesis step. It is clear that for “unimportant” species, one may change their abundance to any reasonable value. However, one should be careful with changing abundances of “important” species, like He, C, N, O, etc, significantly. If this is done, it is recommended to recalculate a full model atmosphere using these new values of abundances. This will not only show whether the previous approach was reasonably accurate, but also the new model may subsequently be used for a fine tuning of abundances.

It should be realized that the calculation of the final synthetic spectrum to be compared to observations involves two steps:

1. producing the net emergent spectrum (radiation flux at the stellar surface) as discussed above; and
2. performing a convolution with rotational and/or instrumental broadening.

Since a star generally rotates, one has to add contributions from all surface elements taking into account the Doppler shift due to the local projected rotational velocity. This procedure is described in detail for instance in Gray (1992). If one assumes a certain a priori given limb-darkening law (i.e. the dependence of specific intensity on  $\mu$ ), one may perform the rotational convolution with the radiation flux; otherwise, the rotational convolution needs a

specification of radiation intensities. The procedure is easy for well-behaving, spherically-symmetric stars. However, one may face complications either because of non-sphericity (for instance that implied by extremely fast rotation), or by departures from surface homogeneity (various starspots, etc.), or for a complicated pattern of velocity fields at the stellar surface (non-radial pulsations, macroturbulence, etc.).

Finally, to be able to compare the predicted spectrum to observations, one has to reproduce numerically a conversion of the incoming stellar spectrum by the spectrograph. In practice, this usually means accounting for a finite spectrum resolution of a spectrograph by convolving the net spectrum with a known instrumental broadening function (usually a Gaussian with a given FWHM). One can also include an instrumental spectrum sensitivity function at this stage. The final result of this step is a predicted spectrum which is directly comparable to the observed one.

### 6.3 Spectrum Fitting

The spectrum fitting is the procedure of finding the model spectrum that fits the observed spectrum best. It may be done by two different ways:

i) a “consecutive model construction” procedure, which consists of a) computing first a small number of initial models for some initial estimates of the basic parameters; b) finding the next estimate of basic model parameters (either by an educated guess, or by using more sophisticated mathematical techniques – for instance the Amoeba optimized search package – Press et al. 1986), and comparing the resulting spectra to observations. The process is repeated until the criteria for a successful match are satisfied. The basic characteristics of this approach is that one does not need any precalculated grid of models; instead, the models are calculated on the way of getting closer and closer to the final model. Obviously, this procedure is efficient only if an effort to generate a model spectrum from scratch is reasonably small.

ii) a “grid-fitting” procedure, which consists in having a precalculated grid of spectra, and finding a model which produces the best fit. One may either find the best-fit model (i.e. one of the models of the grid), or find the best-fit parameters by interpolating in the model grid, assuming that the synthetic spectra corresponding to model parameters in between the grid values may be approximated by an interpolation of tabulated model spectra. If the grid has a sufficiently small step in basic stellar parameters, this procedure is quite satisfactory.

Let us take an example of determining basic stellar parameters for OB stars from observed hydrogen and helium lines. Let us further assume that the mass loss rate is sufficiently low so that all the observed lines originate in the stellar photosphere, i.e. their profiles may be interpreted by means of hydrostatic model atmospheres. Finally, let us assume that we are fitting the observed spectra by means of simple H-He model atmospheres. This means that the grid of spectra depends on five input parameters:  $T_{\text{eff}}$ ,  $\log g$ ,  $Y$  (the

helium abundance),  $v_{\text{tb}}$  (microturbulent velocity), and  $v \sin i$  (projected rotational velocity;  $i$  being the inclination of the rotation axis with respect to the direction to the observer). One may either determine all five of them by a  $\chi^2$ -fitting, or to determine some of them independently, and to fit only a subset of parameters. A typical case for OB main-sequence stars is to determine  $v_{\text{tb}}$  and  $v \sin i$  from metal lines, and to determine the three remaining parameters by a line profile fitting. Sometimes, even  $Y$  may be determined independently (from profiles or equivalent widths of lines of the dominant ion of helium for which the profiles are not so sensitive to  $T_{\text{eff}}$  and  $\log g$ ). One is then left with fitting the observed hydrogen and helium line profiles with only a 2-dimensional grid of spectra which depend only on  $T_{\text{eff}}$  and  $\log g$ .

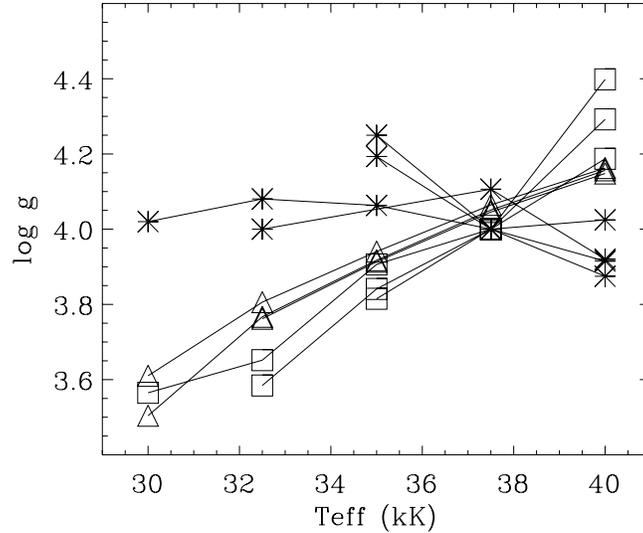
There are two options to perform the actual fitting:

a) fitting both parameters simultaneously; or

b) using the *fit diagrams*. This consists in keeping one parameter fixed (typically  $\log g$ ), and finding such a value of  $T_{\text{eff}}$  which fits the observed profile best. One then goes to the next grid value of  $\log g$ , and repeats the fitting. Every fitted spectral line then defines a curve in the  $T_{\text{eff}}\text{-}\log g$  plane, on which the best-fit values of  $T_{\text{eff}}$  and  $\log g$  are located. Ideally, all curves should intersect in one single point, which then determines the overall best fit values of  $T_{\text{eff}}$  and  $\log g$ . In reality, one usually does not get such a good fit, but at least one should obtain a relatively small region in the  $T_{\text{eff}}\text{-}\log g$  plane where the curves intersect. If one single spectral line defines a significantly different fit curve, it is a strong indication that something on the theoretical or observational level was incorrect.

The fit diagram method is illustrated on the following example: I have constructed a grid of NLTE H-He model atmospheres with effective temperatures between 25000 and 45000 K, in steps of 2500 K, and for  $\log g$  between 3.5 to 4.5, in steps of 0.25. All models have a solar abundance of helium. I will not fit an actual observed spectrum; instead, I will pretend that the “observed” spectrum is the synthetic spectrum computed for a fully metal line-blanketed NLTE model for  $T_{\text{eff}} = 35\,000$  K, and  $\log g = 4$ .

This example will illustrate two features; namely i) what the fit diagrams look like, and ii) what error one makes if the spectrum is fitted by simple H-He model atmospheres instead of by line-blanketed models. The fit diagram for H, He I and He II lines is shown in Fig. 8. A very interesting result is that the H-He models would determine the best fit parameters  $T_{\text{eff}} \approx 38\,000$  K, and  $\log g \approx 4$ . In other words, the H-He models will overestimate the deduced effective temperature, which is not surprising in view of the discussion presented in Sect. 5.5, namely that the local temperature in the H-He models in regions where H and He lines are formed is lower than in the line-blanketed models (no backwarming).



**Fig. 8.** A fit diagram for fitting the H, He I, He II line profiles by means of NLTE H-He model atmospheres. The “observed” spectrum is in fact a synthetic spectrum computed for a fully metal line-blanketed NLTE model for  $T_{\text{eff}} = 35\,000$  K, and  $\log g = 4$ . Squares: hydrogen lines ( $H\beta$  to  $H\delta$ ); triangles: He I lines ( $\lambda\lambda$  4388, 4471, 4922 Å); stars: He II lines ( $\lambda\lambda$  4026, 4200, 4542, 4686 Å).

#### 6.4 Determination of Fundamental Stellar Parameters

Here we will only be concerned with the question how the fundamental stellar parameters are determined from a photospheric analysis, i.e. by analyzing the observed stellar spectrum by means of hydrostatic model atmospheres. An important part of this procedure, which is nevertheless often forgotten, is to verify that the deduced stellar properties are indeed consistent with the assumption of hydrostatic equilibrium.

The fundamental stellar parameters to be determined are the stellar mass,  $M_*$ , radius,  $R_*$ , and luminosity,  $L_*$ . In general, we do not know the distance to the star,  $d$ , so we add this quantity to the list, even if it does not represent an intrinsic stellar property. (There are, obviously, other fundamental stellar parameters, like the chemical composition, rotational velocity, etc. For the purposes of this section, we assume that they are determined independently of the four fundamental parameters listed above.)

The parameters which we determine directly from observations are the

effective temperature,  $T_{\text{eff}}$ , and surface gravity,  $g$ . In addition, we have the measured magnitude,  $m_{\text{obs}}$  that reflects the whole observationally accessible wavelength range. If the flux in the unobservable region is negligible, then this magnitude represent the total, *bolometric*, magnitude,  $m_{\text{bol}}$ . If not, one has to apply the bolometric correction, which follows from the model atmosphere.

In any case, we end up with three “measured” quantities,  $T_{\text{eff}}$ ,  $g$ , and  $m_{\text{bol}}$ , but we have 4 unknown fundamental parameters,  $M_*$ ,  $R_*$ ,  $L_*$ , and  $d$ . The governing relations between them are

$$\sigma T_{\text{eff}}^4 = L_*/(4\pi R_*^2) \quad , \quad (156)$$

$$g = GM_*/R_*^2 \quad , \quad (157)$$

$$L_* = L_* [m_{\text{bol}}(m_{\text{obs}}, T_{\text{eff}}), d] \quad , \quad (158)$$

The last relation expresses the conversion of the observed magnitude to the stellar luminosity.

We thus have three relations for four unknowns. In fact, in some cases the stellar evolution theory may supply an independent additional relation between the fundamental parameters, for instance the mass–radius relation for white dwarfs (Hamada and Salpeter 1961), or the mass–luminosity relation for central stars of planetary nebulae (Paczynski 1971). However, in the general situation we do not have such a relation, and even if we do we may want to check the theoretically predicted relations observationally.

Therefore, from the photospheric analysis only, one cannot derive all four parameters simultaneously. This is easily understood from the physical point of view. A plane-parallel hydrostatic atmosphere is just a thin layer sitting on the top of a spherical star. The only information about a dimension of the underlying star is contained in the surface gravity  $g$  which depends also on the stellar mass. Since the atmosphere is thin, the emergent spectrum does not carry any independent information about the atmospheric extent.

To remove the radius–mass degeneracy, we need either independent geometrical information (knowing the radius or the distance), or an independent knowledge of the mass. The typical situation is that we know the distance  $d$  (the situation will be significantly improved when Hipparchos parallaxes are released); then the other parameters are determined as follows:

1. from known  $m_{\text{obs}}$  and  $d$  (and, possibly,  $T_{\text{eff}}$ ), we determine the absolute bolometric magnitude,  $M_{\text{bol}}$  and, therefore, luminosity,  $L_*$ ;
2. from  $L_*$  and  $T_{\text{eff}}$ , we determine  $R_*$ ;
3. from  $R_*$  and  $g$ , we determine mass  $M_*$

As it turns out, if the mass of early-type O stars is determined in this way (which is called the spectroscopic mass), and if the mass is also determined by comparing the evolutionary tracks and the position of the star in the H-R diagram (the so-called evolutionary mass), one finds a significant discrepancy (e.g. Herrero et al. 1992). The sense of discrepancy is that the spectroscopic

masses are systematically lower than the evolutionary masses. The discrepancy arises either by inaccuracies of the stellar atmospheres theory, or the stellar evolution theory, or, most likely both. From the stellar atmospheres side, there has been a recent progress in understanding the reasons for the discrepancy (e.g. Lanz et al. 1996). However, the problem is not yet solved, and presents a challenge for future research.

## Acknowledgements

It is a pleasure to thank the organizers of the 9th European Astrophysics Doctoral Network Summer School for the invitation to give these lectures. I would also like to thank my colleagues in the audience and my co-lecturers for stimulating discussions. I would like to thank Thierry Lanz, Alex de Koter, Steve Shore, and Tom Brown for their careful reading of the manuscript and offering a number of useful suggestions.

## References

- Allard, F. Hauschildt, P.H. (1995): ApJ **445**, 433  
 Anderson, L. S. (1985): ApJ **298**, 848  
 Anderson, L. S. (1987): in *Numerical Radiative Transfer*, ed. by W. Kalkofen, Cambridge Univ. Press, Cambridge, p. 163  
 Anderson, L. S. (1989): ApJ **339**, 588  
 Anderson, L. S. (1991): in *Stellar Atmospheres: Beyond Classical Models*, ed. by L. Crivellari, I. Hubeny, and D.G.Hummer, NATO ASI Series C 341 , Kluwer, Dordrecht, p. 29  
 Anderson, L. S., Grigsby, J.A. (1991): in *Stellar Atmospheres: Beyond Classical Models*, ed. by L. Crivellari, I. Hubeny, and D.G.Hummer, NATO ASI Series C 341 , Kluwer, Dordrecht, p. 365  
 Auer, L.H. (1987): in *Numerical Radiative Transfer*, ed. by W. Kalkofen, Cambridge Univ. Press, Cambridge, p. 101  
 Auer, L.H. (1991): in *Stellar Atmospheres: Beyond Classical Models*, ed. by L. Crivellari, I. Hubeny, and D.G.Hummer, NATO ASI Series C 341 , Kluwer, Dordrecht, p. 9  
 Auer L.H., Mihalas D. (1969): ApJ **158**, 641  
 Auer, L.H., Mihalas, D. (1970): MNRAS **149**, 60  
 Avrett, E.H., Hummer, D.G. (1965): MNRAS **130**, 295  
 Böhm-Vitense, E. (1989): *Introduction to Stellar Astrophysics II. Stellar Atmospheres*, Cambridge Univ. Press, Cambridge  
 Butler, K., Giddings, J. (1985): in *Newsletter on Analysis of Astronomical Spectra* No. 9, Daresbury Laboratory, Daresbury, England  
 Cannon, C.J. (1973): J. Quant. Spectrosc. Radiat. Transfer **13**, 627  
 Carlsson, M. (1986): *A Computer program for Solving Multi-Level Non-LTe Radiative Transfer Problems in Moving or Static Atmospheres*, Report No. 33, Uppsala Astron. Obs.  
 Castor, J. I., Abbott, D. C., Klein, R. I. (1975): ApJ, **195**, 157

- Castor, J.I., Dykema, P.G., Klein R.I. (1991): in *Stellar Atmospheres: Beyond Classical Models*, ed. by L. Crivellari, I. Hubeny, and D.G.Hummer, NATO ASI Series C 341 , Kluwer, Dordrecht, p. 49.
- Castor, J.I., Dykema, P.G., Klein, R.I. (1992): ApJ, **387**, 561
- Cayrel, R. (1963): C. R. Acad. Sci. Paris **257**, 3309
- Chandrasekhar, S. (1960): *Radiative Transfer*, Dover, New York
- Crivellari, L., Hubeny, I., Hummer D. G. (Eds.) (1991): *Stellar Atmospheres: Beyond Classical Models*, NATO ASI Series C 152 (Dordrecht: Kluwer)
- de Koter, A., Schmutz, W., Lamers, H. J. G. L. M. (1993): A&A , **277**, 561
- Dreizler, S., Werner, K. (1993): A&A, **278**, 199.
- Feautrier, P. (1964): C. R. Acad. Sci. Paris **258**, 3189
- Gabler, R., Gabler, A., Kudritzki, R. P., Puls, J., Pauldrach, A. (1989): A&A, **226**, 162
- Garmany, C.D. (Ed.) (1990): *Properties of Hot Luminous Stars*, PASP Conf. Ser. Vol. 7, Astron. Soc. Pacific, San Francisco
- Gray, D.F. (1992): *Observations and Analysis of Stellar Photospheres*, 2nd ed., Cambridge Univ. Press, Cambridge
- Griem, H.R. (1974): *Spectral Line Broadening by Plasmas*, Acad. Press, New York
- Gustafsson, R., Bell, R.A., Eriksson, K., Nordlund, A. (1975): A&A **42**, 407
- Hamada, T., Salpeter, E.E. (1961): ApJ **134**, 683
- Hamann, W.-R. (1985): A&A **148**, 364
- Heber, U., Jeffery, C.J. (Eds.) (1992): *The Atmospheres of Early-Type Stars*, Lecture Notes in Phys. 401, Springer, Berlin, p. 377
- Herrero, A., Kudritzki, R.P., Vilchez, J.M., Kunze, D., Butler, K., Haser, S. (1992): A&A **261**, 209
- Hillier, D.J. (1990): A&A , **231**, 116
- Hubeny, I. (1985): in *Progress in Stellar Spectral Line Formation Theory*, ed. by J.E. Beckman and L. Crivellari, NATO ASI Series C 152, Reidel, Dordrecht
- Hubeny, I. (1987): A&A **185**, 332
- Hubeny, I. (1988): Comp. Phys. Comm. **52**, 103
- Hubeny, I. (1992): in *The Atmospheres of Early-Type Stars*, ed. by U. Heber and C.J. Jeffery, Lecture Notes in Phys. 401, Springer, Berlin, p. 377
- Hubeny, I., Lanz, T. (1992): A&A **262**, 501
- Hubeny, I., Lanz, T. (1995): ApJ **439**, 875
- Hubeny, I., Lanz, T., Jeffery, C. S. (1994): in *Newsletter on Analysis of Astronomical Spectra* No. 20, Ed. Jeffery C. S., St Andrews Univ., St Andrews, 30
- Hummer, D.G. (1981): J. Quant. Spectrosc. Radiat. Transfer **26**, 187
- Ivanov, V.V. (1973): *Transfer of Radiation in Spectral Lines*, English language translation of Radiative Transfer and the Spectra of Celestial Bodies, transl. by D.G. Hummer, NBS Special Publ. No. 385, National Bureau of Standards, Washington
- Jefferies, J. (1968): *Spectral Line Formation*, Blaisdel, Waltham
- Jeffery, C. S. (1992): in *Newsletter on Analysis of Astronomical Spectra* No. 17, Ed. Jeffery C. S., CCP7, (St Andrews: St Andrews Univ.), 33
- Jeffery, C.J., Heber, U. (Eds.) (1996): *Hydrogen-Deficient Stars*, PASP Conf. Ser. Vol. 96, Astron. Soc. Pacific, San Francisco
- Johnson, H.R., Bernat, A.P., Krupp, B., Bell, R.A. (1977), ApJ **212**, 760

- Kalkofen, W. (Ed.) (1984): *Methods in Radiative Transfer*, Cambridge Univ. Press, Cambridge
- Kalkofen, W. (Ed.) (1987): *Numerical Radiative Transfer*, Cambridge Univ. Press, Cambridge
- Kudritzki, R.P. (1988): *The Atmospheres of Hot Stars: Modern Theory and Observation*, in Radiation in Moving Gaseous Media, ed. by Y. Chmielewski and T. Lanz, 18th Advanced Course, Swiss Soc. of Astron. and Astrophys., Geneva
- Kudritzki, R. P., Hummer D. G. (1990): ARA&A **28**, 303
- Kurucz, R. L. (1970): Smithsonian Astrophys. Obs. Spec. Rep. No. 309
- Kurucz, R. L. (1979): ApJS **40**, 1
- Kurucz, R. L. (1994): Kurucz CD-ROM No. 13, Smithsonian Astrophys. Obs.
- Lanz, T. , Hubeny, I. (1995): ApJ **439**, 905
- Lanz, T., de Koter, A., Hubeny, I., Heap, S. R. (1996): ApJ **465**, 359
- Mihalas, D. (1978): *Stellar Atmospheres*, Freeman, San Francisco
- Mihalas D., Heasley J.N., Auer L.H. (1975): *A Non-LTE model stellar atmospheres computer program*, NCAR-TN/STR-104, NCAR, Boulder
- Pauldrach, A.W.A, Puls, J., Kudritzki, R.P. (1986): A&A, **164**, 86
- Paczynski, B.E. (1971): Acta Astr. **21**, 417
- Olson, G.L., Auer, L.H., Buchler, J.R. (1986): J. Quant. Spectrosc. Radiat. Transfer **35**, 431 (OAB)
- Olson, G.L., Kunasz, P.B. (1987): J. Quant. Spectrosc. Radiat. Transfer **38**, 325
- Oxenius, J. (1986): *Kinetic Theory of Particles and Photons*, Springer-Verlag, Berlin Heidelberg
- Press, W.H., Flannery, B.P., Teukolsky, S.A., Vetterling, W.T. (1986): *Numerical Recipes*, Cambridge Univ. Press, Cambridge
- Rutten, R.J. (1995): *Radiative Transfer in Stellar Atmospheres*, 2nd WWW edition, June 22, 1995, Robert J. Rutten, Utrecht, Netherlands
- Rybicki, G.B. (1971): J. Quant. Spectrosc. Radiat. Transfer **11**, 589
- Rybicki, G.B. (1984): in *Methods in Radiative Transfer*, ed. by W. Kalkofen, Cambridge Univ. Press, Cambridge, p. 21
- Rybicki, G.B. and Hummer, D.G. (1991): A&A, **245**, 171
- Rybicki, G.B. and Hummer, D.G. (1991): A&A, **262**, 209;
- Rybicki, G.B., Lightman, A.P. (1979): *Radiative Processes in Astrophysics*, Wiley-Interscience, New York
- Scharmer, G.B. (1981): ApJ **249**, 720
- Schaerer, D., Schmutz, W. (1994): A&A **228**, 231
- Sellmaier, F., Puls, J., Kudritzki, R.P., Gabler, A., Gabler, R., Voels, S.A, (1993): A&A, **273**, 533
- Shu, F.H. (1991): *The Physics of Astrophysics I. Radiation*, University Science Books, Mill Valley
- Stenflo, J.O. (1994): *Solar Magnetic Fields, Polarized Radiation Diagnostics*, Kluwer, Dordrecht
- Tsuji, T. (1976): Pub. Astron. Soc. Japan **28**, 543
- Werner, K. (1986): A&A **161**, 177
- Werner, K. (1987): in *Numerical Radiative Transfer*, ed. by W. Kalkofen, Cambridge Univ. Press, Cambridge, p. 67
- Werner, K. (1989): A&A **226**, 265
- Werner, K., Husfeld, D. (1985): A&A **148**, 417