A Series of Books in Astronomy and Astrophysics
EDITORS: *Geoffrey Burbidge and Margaret Burbidge*

SECOND EDITION

# *Stellar Atmospheres*

*Dimitri Mihalas*
HIGH ALTITUDE OBSERVATORY
NATIONAL CENTER FOR ATMOSPHERIC RESEARCH

W. H. FREEMAN AND COMPANY
*San Francisco*

# *Contents*

# *Preface*

Since the appearance of the first edition of this book, there has been continuing rapid development of our understanding of stellar atmospheres, and it has been clear to me for some time that a new edition was needed. One of the major motivations for producing a new version of the book at this time is the desire to describe the major advances that have been made—in developing methods to solve the transfer equation in moving media, and in the theory of stellar winds. As was true in the first edition, I have not attempted to cover every possible aspect of the subject, but have again treated a limited number of problems in some depth.

It was clear from the outset that, in view of the great demands made upon the student's time in the now-crowded astrophysics curriculum (resulting from the explosion of our knowledge about the Universe), it was pointless to write a book significantly longer than the first edition. Thus, to add new material, it has been necessary to economize the presentation of the old material, and to omit topics that are of specialized interest or that lie outside the mainstream of the developments of primary importance to the book. In particular, given that today's student is most likely to learn what he knows about radiative transfer in a stellar-atmospheres course, but will be interested in applying it to *other* physical situations, I have purposely shifted the emphasis away from strictly stellar applications, and have developed the transfer theory more generally and completely. I believe that a thorough understanding of the radiative transfer theory presented in this book will equip the student to attack a wide variety of transfer problems, whether in the laboratory, the atmospheres of stars and planets, the interstellar medium, X-ray sources, or quasars. Further, I have added exercises in which the student is asked to fill in missing steps of derivations, or to apply the theory himself to simple examples. In most cases the exercises are quite straightforward

and should require only a few minutes work; but some of the exercises in Chapter 7 require substantial effort and would make good class projects.

Ideally the material in this book should be taught in a course lasting two quarters, covering Chapters 1–7 in the first quarter, and Chapters 8–15 in the second. If an entire year (two semesters) is available, the book should be supplemented with extra material on subjects of interest to the instructor and students, perhaps drawn from problems of solar physics, stellar spectroscopy, pulsating atmospheres, peculiar stars, abundance analyses, or many others. If only one semester is available, I recommend omitting, first, Chapters 4 and 9 (which are more physics than astrophysics); next, Chapters 3 and 10 (which are fairly elementary and may well have been covered in an earlier course); and, finally, if necessary, Chapter 13 (which is not absolutely essential for a basic understanding of line-formation).

In any case, many fascinating subjects will inevitably be omitted, and teacher and student alike may feel frustrated, as I have been in writing the book, that a more complete coverage is not possible. Again and again I have felt like the traveler in Frost's "The Road Not Taken" (**388**, 105)*, in choosing one of two equally fair paths, knowing full well that way would lead on to way, and that I should not return to the other. I only hope that the students will discover for themselves these other charming paths and will spend a pleasant lifetime in their exploration.

It is no longer possible for me to acknowledge fairly the many people who have helped me learn about stellar atmospheres and line-formation, and I shall not try here, beyond offering a sincere thanks to all in whose debt I am. But I would be remiss if I did not specifically thank Lawrence Auer, David Hummer, and George Rybicki, who (as colleagues, critics, teachers, collaborators, and friends) have greatly deepened and enlarged my understanding of the material in this book. Further, I wish to record my great debt to Professor W. W. Morgan of Yerkes Observatory. His encouragement has stimulated much of the work I have done in the past several years, and his wise counsel has greatly enhanced its value. I also thank him for sharing with me a few glimpses of his perception of the nature of scientific method from the lofty point at which he can view it.

I wish in addition to thank the people who have helped with the writing of this book: Barbara Mihalas, for reading and correcting the manuscript and the typescript; Tom Holzer and Richard Klein for reading and commenting upon Chapter 15; and David Hummer and Paul Kunasz for reading the typescript and offering many corrections and suggestions. Thanks also are due to Gordon Newkirk for helping to provide, through his labors as

* A NOTE ABOUT REFERENCES: References are listed serially at the end of the text, and are denoted in the text with boldface numbers—e.g., (**105**). Additional information, such as a page or chapter citation, will be indicated following the reference number—e.g., (**105**, 27) or (**105**, Chap. 4). Citations to two or more references are separated by semicolons—e.g., (**105**, 27; **388**, 105).



Director of H.A.O., the scientific environment in which this book could be written. I also thank Paulina Franz for converting hundreds of pages of my spidery handwriting into smooth typed copy, Kathlyn Auer for preparing the index, and Pat Brewer of W. H. Freeman and Company for her effective and careful supervision of the production process.

Finally, I thank my father, M. D. Mihalas, for his unintentional (but priceless) contribution in teaching me, through the example of his life, the meaning of *αὐτοπεποίθησις* and *φιλοτιμία*.

*Oxford, England* *Dimitri Mihalas*
*October, 1977*

# *Preface*

## TO THE FIRST EDITION

The study of stellar atmospheres is in many ways one of the most interesting and rewarding areas of modern astrophysics. It is not an exaggeration to state that most of what we know about stars, and systems of stars, is derived from an analysis of their radiation, and that this knowledge will be secure only as long as the analytical technique is physically reliable. It is therefore important to have a sound theoretical framework upon which our inferences can be based with confidence.

The field of stellar atmospheres enjoyed a period of rapid growth during the last decade. On the one hand, great improvements were made in the quantity and quality of the observational material. Not only did ground-based observations provide a continuing flow of data, but, in addition, observations from balloons, rockets, and satellites opened broad horizons hitherto completely hidden from us. On the other hand, enormous progress was made in the development of theory. A concerted effort by astronomers and physicists filled in many of the gaps in our understanding of the basic physical processes occurring in stellar atmospheres. The availability of large-capacity, high-speed computing machines stimulated the development of powerful new mathematical techniques and allowed their application to a wide range of cases. Thus, tremendous gains were made in enlarging and clarifying the formal and conceptual basis of the subject.

One of the unfortunate side effects of this period of growth is that practically all of the available textbooks in this field of astronomy are badly out of date. Students and instructors alike must now attempt to digest a large body of scattered literature in order to learn of recent developments. There is, in my mind, a definite need for a new text on the subject, and the present book is an attempt to provide such a text on an introductory graduate level. It is based upon courses I have given to first- and second-year graduate

students at Princeton University, the University of Colorado, and the University of Chicago. It represents what I feel is a minimum background for a student who wishes to understand the literature and to do research in the field. Naturally, it has been necessary to be selective in the material presented. In writing this book, I had in mind the goal of providing a basic synopsis of the theory that can be covered in two quarters, with the hope that the content of the third quarter of the normal academic year will be drawn by the instructor (and the students) from the current literature on topics of special interest to them. Although emphasis is given to the more modern approaches, I have also attempted to give a coherent review of the older methods and results. I feel it is important for students to be familiar with these classical approaches so that they will be aware of the limitations of such approaches and the conclusions based upon them.

It has been tempting to include a wider range of subjects, but I have avoided doing so in the belief that it is more worthwhile for the student to consider a smaller number of topics in depth than attempt to survey the entire field superficially. In this vein, I have purposely limited the comparison of theory with observation to a few of the more crucial and illustrative examples. Moreover, I have restricted most of the theoretical discussion to what may be called the *classical stellar-atmospheres* problem—i.e., atmospheres in hydrostatic, radiative, and steady-state statistical equilibrium. This is ample material for a two-quarter course and is understood well enough to require little speculation. Even within this problem, I have limited the variety of techniques treated. For example, I personally favor using differential equations over using integral equations to solve transfer problems. Thus, although the latter method has enjoyed wide application and good success, particularly in the hands of the Harvard–Smithsonian Astrophysical Observatory group, there is little discussion of it in this book. This omission is not arbitrary, however, but is based upon the view that, since the two methods are mathematically equivalent, discussion of one suffices and, in addition, that the one I have chosen seems to offer more promise in future applications—for example, to situations involving hydrodynamics (wherein lies the real frontier of the subject). On the other hand, in my experience, the physics background of astronomy students is often uneven; I have, therefore, not hesitated to develop those aspects of physical theory that are of special interest to the atmospheres problem. In any case, I hope that users of this book will find it a helpful outline, which they can edit, alter, and enlarge upon as their needs dictate.

*Williams Bay, Wisconsin*
*November 1969*

*Dimitri Mihalas*

# *Stellar Atmospheres*

# 1

# *The Radiation Field*

From quantitative examination of the spectrum of a star, information can be obtained about the frequency distribution of the emergent radiation field. We observe both broad, smooth expanses of *continuum* and *spectrum lines*, where the frequency variation is quite abrupt. The entire spectrum contains an enormous wealth of information, and the primary goal of the theory of stellar atmospheres is to develop methods that can recover this information. To this end we must be able to describe the flow of energy through the outermost layers of a star, and to predict the observable characteristics of the emergent radiation. We apply known physical laws that specify the interaction of radiation with stellar material, and derive mathematical models from which we compute theoretical estimates of observables. We then compare theory and observation, and attempt to infer the physical conditions in stellar atmospheres. Such analyses can provide information about the structure of the envelope (important as a boundary condition for studies of stellar structure), modes of energy transport in the atmosphere, chemical abundances, rates of mass loss, and calibrations for converting observational parameters (e.g., $M_v$ and $B - V$) into theoretically interpretable numbers (luminosity and temperature). By studying large numbers of stars we can

establish relations of, say, chemical composition to stellar distributions, kinematics, and dynamics; this information provides clues in developing an understanding of the structure and dynamics of the Galaxy as a whole.

The program outlined above is ambitious, and it is not an easy one to carry out successfully. The observational data are often difficult to acquire, have limited precision, and are the results of very complicated physical structures. Often our physical theories are only primitive, and yet even these may lead to extremely complicated mathematical systems. But the key issue is that the information we deduce from stellar spectra will be a close approximation to reality only if the underlying physical theory is sound and comprehensive. We must, therefore, devote considerable attention to the development of an approach that correctly includes the essential physics.

In this chapter we introduce the basic definitions required to characterize the radiation field itself. The radiation field is treated from three points of view—using macroscopic, electromagnetic, and quantum descriptions. Each of these approaches yields useful information and, taken together, they provide a full picture of the nature of the field. We ignore polarization, but carry along an assumed time-dependence so that in later work we can derive equations of radiation hydrodynamics. In subsequent chapters we shall consider how the radiation interacts with the material and is transported through the atmosphere (Chapter 2), and shall write down detailed descriptions of the atomic parameters that specify the absorptivity of the material (Chapter 4) and the mechanisms that determine the distribution of atoms over available bound and free states (Chapter 5). After consideration of the grey problem, which supplies an ideal testing ground of methods and shows clearly the overall approach used (Chapter 3) and development of general mathematical techniques for solving transfer equations (Chapter 6), we discuss the central problem of the book: the construction of model atmospheres (Chapter 7). We then examine the physics of line formation for a given (static) model (Chapters 8–13), and the methods used to infer chemical abundances in and physical characteristics of stellar atmospheres. Radiative transfer in moving atmospheres is then analyzed (Chapter 14) and, finally, all of the preceding developments are applied in a discussion of stellar winds (Chapter 15).

## 1-1 The Specific Intensity

### MACROSCOPIC DEFINITION

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 amount of



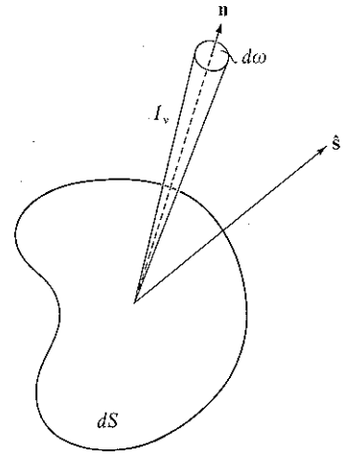



FIGURE 1-1
Pencil of radiation used to define specific intensity. The vector $\mathbf{n}$ is the direction of propagation of the radiation, while $\hat{\mathbf{s}}$ is the unit vector perpendicular to the element of area $dS$.

energy transported by radiation of frequencies $(\nu, \nu + d\nu)$ across an element of area $dS$ into a solid angle $d\omega$ in a time interval $dt$ is

$$\delta\mathscr{E} = I(\mathbf{r}, \mathbf{n}, \nu, t)\, dS \cos\theta\, d\omega\, d\nu\, dt \tag{1-1}$$

where $\theta$ is the angle between the direction of the beam and the normal to the surface (i.e., $dS \cos\theta = \mathbf{n} \cdot \mathbf{dS}$); see Figure 1-1. The dimensions of $I$ are ergs $\mathrm{cm}^{-2}\ \mathrm{sec}^{-1}\ \mathrm{hz}^{-1}\ \mathrm{sr}^{-1}$. As it has just been defined, the specific intensity provides a complete description of the radiation field from a macroscopic point of view.

In this book consideration will be given only to one-dimensional problems in planar or spherical geometry; that is, the atmosphere will be regarded as composed of either homogeneous plane layers or homogeneous spherical shells. In planar geometries we employ Cartesian $(x, y, z)$ coordinates with planes of constant $z$ being the homogeneous layers; we can then ignore the $(x, y)$ dependence of all variables, as well as derivatives with respect to $x$ and $y$. It is convenient to introduce polar and azimuthal angles $(\theta, \phi)$ to specify $\mathbf{n}$; we then have $\mathbf{n} \cdot \mathbf{k} = \cos\theta$, $\mathbf{n} \cdot \mathbf{i} = \sin\theta \cos\phi$, $\mathbf{n} \cdot \mathbf{j} = \sin\theta \sin\phi$. For one-dimensional planar geometry $I$ will clearly be independent of $\phi$; hence we can write $I = I(z, \theta, \nu, t)$; $z$ is measured as positive upward in the atmosphere (opposite to the direction of gravity). In spherical geometry spatial location is specified by $(r, \Theta, \Phi)$; but for spherical *symmetry*, $I$ will depend upon $r$ only. The direction of the radiation can be specified in terms of the polar and azimuthal angles $(\theta, \phi)$, now measured with respect to a unit vector $\hat{\mathbf{r}}$ in the radial direction. Spherical symmetry again implies azimuthal invariance, and we can now write $I = I(r, \theta, \nu, t)$. We shall often replace the variable $\theta$ with $\mu \equiv \cos\theta$.

*Exercise 1-1*: By use of Snell's law, $n_1(\nu) \sin \theta_1 = n_2(\nu) \sin \theta_2$, in the calculation of the energy passing through a unit area on the interface between two dispersive media with differing indices of refraction, show that $I_\nu n_\nu^{-2}$ is a constant.

### PHOTON DISTRIBUTION FUNCTION

The radiation field can also be described in terms of a *photon distribution function* $f_R$ which is defined such that $f_R(\mathbf{r}, \mathbf{n}, \nu, t)\, d\omega\, d\nu$ is the number of photons per unit volume at location $\mathbf{r}$ and time $t$, with frequencies on the range $(\nu, \nu + d\nu)$, propagating with velocity $c$ in direction $\mathbf{n}$ into a solid angle $d\omega$. Each photon has an energy $h\nu$. The number of photons crossing an element $\mathbf{dS}$ in time $dt$ is $f_R(c\, dt)(\mathbf{n} \cdot \mathbf{dS})(d\omega\, d\nu)$, so that the energy transported is $\delta\mathscr{E} = (ch\nu) f_R\, dS \cos\theta\, d\omega\, d\nu\, dt$; comparison of this expression with equation (1-1) shows that

$$I(\mathbf{r}, \mathbf{n}, \nu, t) \equiv (ch\nu) f_R(\mathbf{r}, \mathbf{n}, \nu, t) \tag{1-2}$$

### INVARIANCE PROPERTIES

An important property of the specific intensity is that it has been defined in such a way as to be independent of the distance between the source and the observer if there are no sources or sinks of radiation along the line of sight. Thus, consider that pencil of rays which passes through both the element of area $dS$ at point $P$ *and* the element $dS'$ at $P'$ (see Figure 1-2). Then the amount of energy $\delta\mathscr{E}$ passing through both areas is

$$\delta\mathscr{E} = I_\nu\, dS \cos\theta\, d\omega\, d\nu\, dt = \delta\mathscr{E}' = I'_\nu\, dS' \cos\theta'\, d\omega'\, d\nu\, dt \tag{1-3}$$

where $d\omega$ is the solid angle subtended by $dS'$ as seen from $P$, and $d\omega'$ is the solid angle subtended by $dS$ as seen from $P'$. From Figure 1-2 we see that $d\omega = r^{-2}\, dS' \cos\theta'$ while $d\omega' = r^{-2}\, dS \cos\theta$, where $r$ is the distance from

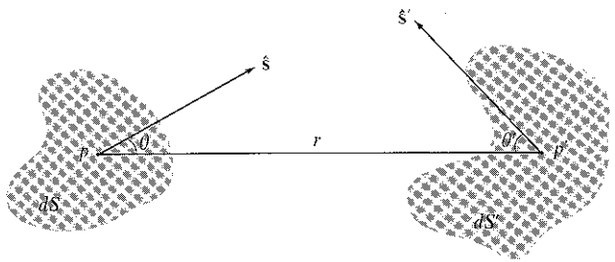



FIGURE 1-2
Geometry used in proof of invariance of specific intensity. The points $P$ and $P'$ are separated by a distance $r$. Area $dS$ subtends a solid angle $d\omega'$ at $P'$, and the area $dS'$ subtends $d\omega$ at $P$; $\hat{\mathbf{s}}$ and $\hat{\mathbf{s}}'$ are unit vectors normal to $dS$ and $dS'$.



$P$ to $P'$. Thus it immediately follows from equation (1-3) that $I_\nu \equiv I'_\nu$. Note also that equation (1-3) implies that the energy received per unit area falls off as the inverse square of the distance between $P$ and $P'$.

### OBSERVATIONAL SIGNIFICANCE

The spatial invariance of the specific intensity implies that the actual value of $I$ at the source can be obtained from measurements of the amount of energy falling, in a given time, within a specified frequency band, onto a receiver of known collecting area (and detection efficiency) from a source subtending a definite solid angle. The requirement that $d\omega$ must be specified limits the determination of $I$ to sources that are *spatially resolved*—e.g., nebulae, galaxies, the sun, planets, etc.

In particular, for the sun, the radiation at a given point emerges at a known angle relative to the local normal (in a one-dimensional model); hence measurement of the center-to-limb variation of the radiation allows us to determine the angular variation of $I$. Note that we do not, in general, see to the same depth in the atmosphere along all rays; hence we do not obtain the angular variation of $I$ at some definite position ($z$) inside the atmosphere, but rather at some point $\mathbf{r}_{\text{obs}}$ *outside* the atmosphere.

*Exercise 1-2*: The angular diameter of the sun is 30′. Suppose that atmospheric seeing effects limit resolution to 1″; show that this sets a lower bound on the $\mu$ for which we can infer $I(\mu)$ accurately, and determine this $\mu_{\min}$.

## 1-2 Mean Intensity and Energy Density

### MACROSCOPIC DESCRIPTION

In both the physical and the mathematical description of a radiation field it is useful to employ various angular averages, or *moments*. Thus we define the *mean intensity* to be the straight average (zero-order moment) of the specific intensity over all solid angles, i.e.,

$$J(\mathbf{r}, \nu, t) = (4\pi)^{-1} \oint I(\mathbf{r}, \mathbf{n}, \nu, t)\, d\omega \tag{1-4}$$

The mean intensity has dimensions ergs $\text{cm}^{-2}\ \text{sec}^{-1}\ \text{hz}^{-1}$. The element of solid angle $d\omega$ is given by $d\omega = \sin\theta\, d\theta\, d\phi = -d\mu\, d\phi$. If we consider one-dimensional atmospheres, $I$ is independent of $\phi$, hence

$$J(z, \nu, t) = (4\pi)^{-1} \int_0^{2\pi} d\phi \int_{-1}^{1} d\mu\, I(z, \mu, \nu, t) = \frac{1}{2} \int_{-1}^{1} I(z, \mu, \nu, t)\, d\mu \tag{1-5}$$

The same result applies in spherical geometry with $z$ replaced by $r$.

To calculate the *energy density* in the radiation field on the frequency range $(\nu, \nu + d\nu)$, consider a small volume $V$ through which energy flows from all solid angles. The amount flowing from a particular solid angle $d\omega$ through an element of surface area $dS$ of this volume is

$$\delta\mathscr{E} = I(\mathbf{r}, \mathbf{n}, \nu, t)(dS \cos\theta)\, d\omega\, d\nu\, dt$$

Now consider only those photons in flight across $V$; if the path length across $V$ is $l$, then the time they will be contained within $V$ is $dt = l/c$. Further, $l\, dS \cos\theta = dV$, the differential element of $V$ through which they sweep. Hence the energy in $dV$ coming from $d\omega$ is $\delta\mathscr{E} = c^{-1} I(\mathbf{r}, \mathbf{n}, \nu, t)\, d\omega\, d\nu\, dV$; by integrating over all solid angles and over the entire volume, we find the total energy contained in $V$, namely:

$$\mathscr{E}(\mathbf{r}, \nu, t)\, d\nu = c^{-1}\left[\int_V dV \oint d\omega\, I(\mathbf{r}, \mathbf{n}, \nu, t)\right] d\nu \tag{1-6}$$

But if we pass to the limit of *infinitesimal* $V$, $I$ becomes independent of position in $V$, and the integrations can be carried out separately. The *monochromatic energy density*, $E_R(\mathbf{r}, \nu, t) \equiv \mathscr{E}(\mathbf{r}, \nu, t)/V$ is thus

$$E_R(\mathbf{r}, \nu, t) = c^{-1} \oint I(\mathbf{r}, \mathbf{n}, \nu, t)\, d\omega = (4\pi/c) J(\mathbf{r}, \nu, t) \tag{1-7}$$

$E_R$ has dimensions of ergs $\text{cm}^{-3}$ $\text{hz}^{-1}$. The *total energy density* (dimensions: ergs $\text{cm}^{-3}$) is found by integrating over all frequencies:

$$E_R(\mathbf{r}, t) = \int_0^\infty E_R(\mathbf{r}, \nu, t)\, d\nu = (4\pi/c) \int_0^\infty J(\mathbf{r}, \nu, t)\, d\nu \equiv (4\pi/c) J(\mathbf{r}, t) \tag{1-8}$$

PHOTON PICTURE

It is easy to show that the results derived above are consistent with the photon picture of the radiation field. By definition, $f_R(\mathbf{r}, \mathbf{n}, \nu, t)$ is the number of photons, per unit volume, of energy $h\nu$ propagating in direction $\mathbf{n}$ into intervals $d\nu\, d\omega$. The energy density clearly is just this number, multiplied by the energy per photon, summed over all solid angles: i.e.,

$$E_R(\mathbf{r}, \nu, t) = h\nu \oint f_R(\mathbf{r}, \mathbf{n}, \nu, t)\, d\omega \tag{1-9}$$

But from equation (1-2), $h\nu f_R = c^{-1} I$, hence equation (1-9) is seen to be identical with equation (1-7).

EQUILIBRIUM VALUE

In *thermal equilibrium* the radiation field inside an adiabatic enclosure is uniform, isotropic, time-independent, and has a frequency distribution



given by the *Planck function* $B_\nu(T) = (2h\nu^3/c^2)(e^{h\nu/kT} - 1)^{-1}$ [see (**520**), (**392**, 365)]. Thus, in thermal equilibrium the monochromatic energy density is $E_R^*(\nu) = (4\pi/c) B_\nu(T)$, and the total energy density is given by *Stefan's law*:

$$E_R^* = (8\pi h/c^3) \int_0^\infty (e^{h\nu/kT} - 1)^{-1} \nu^3\, d\nu = a_R T^4 \tag{1-10}$$

where $a_R = 8\pi^5 k^4/(15 c^3 h^3)$. Here, as elsewhere in this book, we denote a quantity computed from thermodynamic equilibrium relations with an asterisk.

*Exercise 1-3*: Derive Stefan's law by substituting $x \equiv h\nu/kT$, and expanding $(e^x - 1)^{-1} = e^{-x}(1 - e^{-x})^{-1}$ as a power series in $e^{-x}$. The sum obtained from the term-by-term integration is related to the Riemann zeta-function [see (**4**, 807)].

Stefan's law is valid in the interior of a star, and in the deeper layers of stellar atmospheres, where thermal gradients over a photon mean-free-path are extremely small, and the radiation becomes isotropic and thermalizes to its equilibrium value. At the surface, the radiation field becomes very anisotropic and has a markedly non-Planckian frequency distribution, as a result of steep temperature gradients and the existence of an open boundary through which photons escape into interstellar space; here Stefan's law becomes invalid.

ELECTROMAGNETIC DESCRIPTION

Electromagnetic theory provides an alternative description of the radiation field; we shall show how a one-to-one correspondence can be made between the macroscopic and electromagnetic descriptions of the radiation field. The electromagnetic field is specified by *Maxwell's equations* [see, e.g., (**331**, Chap. 6)] which, in Gaussian units, are

$$\nabla \cdot \mathbf{D} = 4\pi\rho \tag{1-11a}$$

$$\nabla \cdot \mathbf{B} = 0 \tag{1-11b}$$

$$(\nabla \times \mathbf{E}) + c^{-1}(\partial\mathbf{B}/\partial t) = 0 \tag{1-11c}$$

and

$$(\nabla \times \mathbf{H}) - c^{-1}(\partial\mathbf{D}/\partial t) = (4\pi/c)\mathbf{j} \tag{1-11d}$$

The *electric field* $\mathbf{E}$ is related to the *electric displacement* $\mathbf{D}$ in terms of the *permittivity* $\varepsilon$, namely $\mathbf{D} = \varepsilon\mathbf{E}$. Similarly, the *magnetic induction* $\mathbf{B}$ can be expressed in terms of the *magnetic field* $\mathbf{H}$ and the *permeability* $\mu$ by the relation $\mathbf{B} = \mu\mathbf{H}$. For vacuum, $\varepsilon = \mu = 1$. In equations (1-11), $\rho$ is the *charge density* and $\mathbf{j}$ is the *current density* $\mathbf{j} = \rho\mathbf{v}$ associated with charges moving with velocity $\mathbf{v}$. The electric field and magnetic induction can be derived from a scalar potential $\phi$ and a vector potential $\mathbf{A}$, which are defined

such that

$$\mathbf{B} = \nabla \times \mathbf{A} \tag{1-12a}$$

and

$$\mathbf{E} = -\nabla\phi - c^{-1}(\partial\mathbf{A}/\partial t) \tag{1-12b}$$

Equation (1-12a) satisfies equation (1-11b), while (1-12b) satisfies (1-11c). Because **B** is defined as the *curl* of **A**, the *divergence* of **A** may be specified arbitrarily; one of the most convenient choices is to impose the *Lorentz condition*

$$\nabla \cdot \mathbf{A} = -c^{-1}(\partial\phi/\partial t) \tag{1-13}$$

With this choice, Maxwell's equations can be reduced to

$$\nabla^2\phi - c^{-2}(\partial^2\phi/\partial t^2) = -4\pi\rho \tag{1-14a}$$

and

$$\nabla^2\mathbf{A} - c^{-2}(\partial^2\mathbf{A}/\partial t^2) = -(4\pi/c)\mathbf{j} \tag{1-14b}$$

The solutions of these equations can be written as [cf. (**331**, Chap. 6), (**494**, Chap. 19)],

$$\phi(\mathbf{r}, t) = \int \frac{\rho(\mathbf{r}', t')}{|\mathbf{r} - \mathbf{r}'|}\, d^3r' \tag{1-15a}$$

and

$$\mathbf{A}(\mathbf{r}, t) = \frac{1}{c}\int \frac{\rho(\mathbf{r}', t')\mathbf{v}(\mathbf{r}', t')}{|\mathbf{r} - \mathbf{r}'|}\, d^3r' \tag{1-15b}$$

where, as indicated, $\rho$ and $\mathbf{v}$ at $\mathbf{r}'$ are evaluated at the *retarded time* $t' = t - c^{-1}|\mathbf{r}' - \mathbf{r}|$ which takes into account the finite speed of propagation of electromagnetic waves.

One of the most important solutions of Maxwell's equations is that for *monochromatic plane waves in vacuum*, propagating in direction $\mathbf{n}_0$ with velocity $c$:

$$\mathbf{E}(\mathbf{r}, t) = \mathbf{E}_0 \cos[2\pi(k\mathbf{n}_0 \cdot \mathbf{r} - \nu t)] \tag{1-16a}$$

and

$$\mathbf{H}(\mathbf{r}, t) = \mathbf{H}_0 \cos[2\pi(k\mathbf{n}_0 \cdot \mathbf{r} - \nu t)] \tag{1-16b}$$

where $k = \lambda^{-1} = c^{-1}\nu$. The vectors $(\mathbf{E}_0, \mathbf{H}_0, \mathbf{n}_0)$ form an orthogonal triad with $\mathbf{H}_0 = \mathbf{n}_0 \times \mathbf{E}_0$, so it follows that $|\mathbf{H}_0| = |\mathbf{E}_0|$. The result obtained from electromagnetic theory for the instantaneous energy density $W(t)$ in the field is

$$W(t) = (\mathbf{E} \cdot \mathbf{D} + \mathbf{B} \cdot \mathbf{H})/8\pi \tag{1-17}$$

Averaging in time over a cycle introduces a factor of $\langle\cos^2 \omega t\rangle_T = \frac{1}{2}$, and using the relations $|E_0| = |H_0|$ and $\mu = \varepsilon = 1$ (for vacuum), equation (1-17) reduces to $W = \langle W(t)\rangle_T = E_0^2/8\pi$. In terms of the macroscopic picture, a monochromatic plane wave propagating in direction $\mathbf{n}_0$ [specified by angles $(\theta_0, \phi_0)$] has a specific intensity $I(\mu, \phi) = I_0\, \delta(\mu - \mu_0)\, \delta(\phi - \phi_0)$ where $\delta$ denotes the usual Dirac function. Substitution into equation (1-7) yields the energy density $E_R = c^{-1}I_0$, a result that is intuitively obvious for a plane wave propagating with velocity $c$. Therefore we obtain a correspondence between the two descriptions by making the identification



$$I_0 = cE_0^2/8\pi \tag{1-18}$$

It will be shown below that this choice yields consistent relations between the Poynting vector and Maxwell stress tensor and their macroscopic counterparts. The results derived here apply, strictly, only to a monochromatic plane wave, but are easily generalized to fields having arbitrary angle and frequency distributions by summing over suitably chosen elementary plane waves.

## 1-3 The Flux

### MACROSCOPIC DESCRIPTION

We define the *flux* of radiation $\mathscr{F}(\mathbf{r}, \nu, t)$ as a vector quantity such that $\mathscr{F} \cdot \mathbf{dS}$ gives the *net rate of radiant energy flow* across the arbitrarily oriented surface **dS** per unit time and frequency interval. Noting that $\mathbf{n} \cdot \mathbf{dS} = dS \cos\theta$, where $\theta$ is the angle between the direction of propagation **n** and the normal to $dS$, we immediately recognize that the flux can be derived from the specific intensity via equation (1-1), for $\delta\mathscr{E}$ as written there is, in fact, nothing more than the contribution of the pencil of radiation moving in direction **n** to the net energy flux. Thus we merely sum over all solid angles and obtain

$$\mathscr{F}(\mathbf{r}, \nu, t) = \oint I(\mathbf{r}, \mathbf{n}, \nu, t)\mathbf{n}\, d\omega \tag{1-19}$$

The flux has dimensions: ergs $\text{cm}^{-2}$ $\text{sec}^{-1}$ $\text{hz}^{-1}$. Note that $\mathscr{F}$ is the first moment of the radiation field with respect to angle.

In cartesian coordinates we have

$$(\mathscr{F}_x, \mathscr{F}_y, \mathscr{F}_z) = \left(\oint In_x\, d\omega, \oint In_y\, d\omega, \oint In_z\, d\omega\right) \tag{1-20}$$

where $d\omega = -d\mu\, d\phi$, $n_x = (1 - \mu^2)^{\frac{1}{2}} \cos\phi$, $n_y = (1 - \mu^2)^{\frac{1}{2}} \sin\phi$, $n_z = \mu$. If the radiation field is symmetric with respect to an axis, it follows that there will be a ray-by-ray cancellation in the net energy transport across a surface oriented perpendicular to that axis, and that the net flux is identically zero across this surface. In particular, for a planar atmosphere homogeneous in $x$ and $y$, only $\mathscr{F}_z$ can be nonzero; we shall therefore require only this component of the flux, and shall refer to it as "the" flux, as if it were a scalar,

and write

$$\mathscr{F}(z, \nu, t) = 2\pi \int_{-1}^{1} I(z, \mu, \nu, t)\mu \, d\mu \tag{1-21}$$

*Exercise 1-4*: (a) Show that $\mathscr{F}_x$ and $\mathscr{F}_y$ vanish in an atmosphere with azimuthal ($\phi$) independence of $I$. (b) Show that in a spherically symmetric atmosphere only $\mathscr{F}_r$ is nonzero and is given by equation (1-21) with $z$ replaced by $r$. (c) Evaluate $\mathscr{F}$ for $I(\mu) = \sum I_n \mu^n$; show that only the odd-order terms contribute to $\mathscr{F}$.

In astrophysical work it is customary to absorb the factor of $\pi$ appearing in equation (1-21), and to write the *astrophysical flux* as $F(z, \nu, t) \equiv \pi^{-1}\mathscr{F}(z, \nu, t)$. Further, regarding the flux as one of a sequence of moments with respect to $\mu$, one may define the *Eddington flux*

$$H(z, \nu, t) \equiv (4\pi)^{-1}\mathscr{F}(z, \nu, t) = \frac{1}{2}\int_{-1}^{1} I(z, \mu, t)\mu \, d\mu \tag{1-22}$$

which is in a form similar to equation (1-5) for the mean intensity.

### PHOTON ENERGY FLUX

The same results for the energy flux may be obtained from the description of the radiation field in terms of photons. The *net number* of photons passing, with velocity $c$, through a unit surface oriented at angle $\theta$ to the beam, per unit time, is clearly

$$N(\mathbf{r}, \nu, t) = c \oint f_R(\mathbf{r}, \mathbf{n}, \nu, t) \cos\theta \, d\omega \tag{1-23}$$

Each photon has energy $h\nu$, so the net energy transport must be

$$\mathscr{F}(\mathbf{r}, \nu, t) = (ch\nu) \oint f_R(\mathbf{r}, \mathbf{n}, \nu, t)\mathbf{n} \, d\omega \tag{1-24}$$

In view of equation (1-2), equation (1-24) is obviously identical to equation (1-19).

Furthermore, photons of energy $h\nu$ propagating in direction $\mathbf{n}$ have momentum $h\nu\mathbf{n}/c$. Thus it is clear that $c^{-1}\mathscr{F} \cdot \mathbf{dS}\, dt$ gives the net momentum transport across the surface $dS$ in time $dt$, by particles moving with velocity $c$. It therefore follows that the *momentum density* associated with the radiation field is $\mathbf{G}_R = c^{-2}\mathscr{F}$; we shall find further significance of this result in §2-3 and shall use it in §14-3.

*Exercise 1-5*: Verify the assertion that $c^{-2}\mathscr{F}$ represents a momentum density; check units for consistency.



### THE POYNTING VECTOR

In electromagnetic theory, the energy flux in the field is given by the *Poynting vector*

$$\mathbf{S} = (c/4\pi)(\mathbf{E} \times \mathbf{H}) \tag{1-25}$$

Considering a plane wave as in §1-2, the average power over a cycle is $\langle \mathbf{S} \rangle_T = c\langle \mathbf{E} \times \mathbf{H} \rangle_T/4\pi = (c\langle E^2 \rangle_T \mathbf{n}_0)/4\pi = (cE_0{}^2)\mathbf{n}_0/8\pi$. On the other hand, in terms of macroscopic quantities, the flux associated with a plane wave is

$$\mathscr{F} = \oint I\mathbf{n} \, d\omega = \oint I_0 \, \delta(\mathbf{n} - \mathbf{n}_0)\mathbf{n} \, d\omega = I_0\mathbf{n}_0 \tag{1-26}$$

Now using equation (1-18), it is clear that $\mathscr{F}$ defined by equation (1-26) is identical to $\langle \mathbf{S} \rangle_T$. Again, this result can be generalized to arbitrary angle and frequency distributions of the radiation field.

### OBSERVATIONAL SIGNIFICANCE

The energy received from a star by a distant observer can be related directly to the flux $\mathscr{F}_\nu$ emitted at the stellar surface. Assume that the distance $D$ between star and observer is very much larger than the stellar radius $r_*$, so that all rays from star to observer may be considered to be parallel. The energy received, per unit area normal to the line of sight, from a differential area on the star is $df_\nu = I_\nu \, d\omega$ where $d\omega$ is the solid angle subtended by the area, and $I_\nu$ is the specific intensity emergent at the stellar surface. Considering the geometry shown in Figure 1-3 we see that $r = r_* \sin\theta$ so that the area of

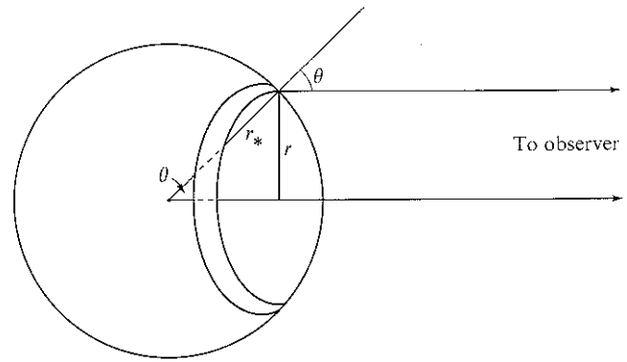



FIGURE 1-3
Geometry of measurement of stellar flux. The annulus on the surface of the star has an area $dS = 2\pi r \, dr = 2\pi r_*^2 \sin\theta \cos\theta \, d\theta$ normal to the line of sight; this area subtends a solid angle $d\omega = dS/D^2$ as seen by the observer.

a differential annulus on the disk is $dS = 2\pi r\, dr = 2\pi r_*^2 \mu\, d\mu$, and $d\omega = 2\pi(r_*/D)^2 \mu\, d\mu$. The radiation emitted from this annulus in the direction of the observer emerged at angle $\theta$ relative to the normal; hence the appropriate value of the specific intensity is $I(r_*, \mu, \nu)$. Integrating over the disk, we find

$$\mathscr{F}_\nu = 2\pi(r_*/D)^2 \int_0^1 I(r_*, \mu, \nu)\mu\, d\mu = (r_*/D)^2 \mathscr{F}(r_*, \nu) = \frac{1}{4}\alpha_*^2 \mathscr{F}(r_*, \nu) \quad (1\text{-}27)$$

where $\alpha_*$ is the *angular diameter* of the star. [In the above calculation we have assumed there is no radiation incident upon the surface of the star; i.e., $I(r_*, -\mu, \nu) \equiv 0$.] For *unresolved* objects (e.g., stars), we can measure only the flux. The energy received falls off as the inverse square of the distance (because the solid angle subtended by the disk varies as $D^{-2}$). If the angular diameter is known, then the absolute energy flux measured at the earth can be converted to the absolute flux at the star.

*Exercise 1-6*: Show that the flux emergent from a small aperature in an adiabatic enclosure (blackbody) is $\mathscr{F}_{\nu B}(\nu) = \pi B_\nu(T)$. Show that the *integrated* flux is $\mathscr{F}_B = \sigma_R T^4$ where $\sigma_R = (c/4)a_R = 2\pi^5 k^4/(15h^3c^2) = 5.67 \times 10^{-5}$ erg cm$^{-2}$ sec$^{-1}$ deg$^{-4}$ is the *Stefan-Boltzmann constant*.

## 1-4 The Radiation Pressure Tensor

### MACROSCOPIC DESCRIPTION AND THE PHOTON MOMENTUM FLUX

The mean intensity and flux are the scalar and vector quantities given by the zero and first angular moments of the specific intensity against the direction cosines between the direction of propagation and an orthogonal triad. The *second* moment yields a tensor quantity, that we shall identify as the *radiation pressure tensor* (or *radiation stress tensor*), namely

$$\mathsf{P}(r, \nu, t) = c^{-1} \oint I(\mathbf{r}, \mathbf{n}, \nu, t)\mathbf{nn}\, d\omega \quad (1\text{-}28)$$

or, in component form,

$$P_{ij}(r, \nu, t) = c^{-1} \oint I(\mathbf{r}, \mathbf{n}, \nu, t) n_i n_j\, d\omega \quad (1\text{-}29)$$

The dimensions of P are ergs cm$^{-3}$ hz$^{-1}$. It is obvious that P is *symmetric*; i.e., $P_{ij} = P_{ji}$.

The physical interpretation of P follows directly from the description of the radiation field in terms of photons. Thus, using equation (1-2) to replace the specific intensity with the photon distribution function $f_R$, we see that

$$P_{ij}(\mathbf{r}, \nu, t) = \oint [f_R(\mathbf{r}, \mathbf{n}, \nu, t)cn_i](h\nu n_j/c)\, d\omega \quad (1\text{-}30)$$

The above expression clearly gives the net flux of momentum, in the $j$-direction, per unit time, from radiation of frequency $\nu$, through a unit area oriented perpendicular to the $i$-direction; this is precisely the definition of *pressure* in any fluid, and hence justifies the term "radiation pressure".

The average of the diagonal components of P may be used to define a *mean radiation pressure*:

$$\bar{P} = \frac{1}{3}(P_{xx} + P_{yy} + P_{zz}) \quad (1\text{-}31)$$

But $(n_x^2 + n_y^2 + n_z^2) \equiv 1$ for any unit vector $\mathbf{n}$, hence in general

$$\bar{P}(\mathbf{r}, \nu, t) = (3c)^{-1} \oint I(\mathbf{r}, \mathbf{n}, \nu, t)\, d\omega = \frac{1}{3} E_R(\mathbf{r}, \nu, t) \quad (1\text{-}32)$$

However, it must be emphasized that despite the generality of this result, $\bar{P}$ does *not* give the actual radiation pressure unless the radiation field happens to be *isotropic*. In general the radiation field in stellar atmospheres is far from isotropic, and ordinarily the numerical factor relating $p_R$ (a scalar parameter that can be used to calculate radiation forces) and the energy density $E_R$ *exceeds* $\frac{1}{3}$ (see below).

### RELATION OF THE PRESSURE TENSOR TO VOLUME FORCES

Let us now examine the relation of the radiation pressure tensor to volume forces exerted by the radiation field. Consider an element of area $dS$; the flow, per unit time, of the $i$-component of momentum in the radiation field across this element is $\sum_j P_{ij} n_j\, dS$, where the $n_j$'s are the direction cosines of the normal to $dS$. Now integrating over a closed surface $S$, and applying the divergence theorem, we find

$$\oint_S \sum_j P_{ij} n_j\, dS = \int_V \sum_j (\partial P_{ij}/\partial x_j)\, dV = \int_V (\nabla \cdot \mathsf{P})_i\, dV \quad (1\text{-}33)$$

where $V$ is the volume enclosed by $S$. The integral on the left gives the net flow, per unit time, of the $i$-component of momentum out of the volume through the surface $S$; thus from the integral on the right we see that $(\nabla \cdot \mathsf{P})_i$ must be the rate at which the $i$-component of the momentum density in the field decreases; i.e., $c^{-2}(\partial \mathscr{F}/\partial t)_i$.

Hence *for the radiation field alone* (i.e., in the absence of absorbing or emitting material) we have

$$(\partial \mathbf{G}_R/\partial t) = c^{-2}[\partial \mathcal{F}(\mathbf{r}, \nu, t)/\partial t] = -\nabla \cdot \mathsf{P}(\mathbf{r}, \nu, t) \tag{1-34}$$

Equation (1-34) is essentially identical to the usual momentum equation of hydrodynamics for an ideal fluid with no applied forces (cf. §15-1). We shall generalize this result to include interactions with material in §2-3.

### THE MAXWELL STRESS TENSOR

In electromagnetic theory, the stress in the field is described by the *Maxwell stress tensor*, which is defined such that

$$(\partial \mathbf{G}_{\text{em}}/\partial t) = \nabla \cdot \mathsf{T}^M \tag{1-35}$$

Here $\mathbf{G}_{\text{em}}$ is the momentum density associated with the electromagnetic field. The components of $\mathsf{T}^M$ are:

$$T_{ij}{}^M = \left[E_iE_j + H_iH_j - \frac{1}{2}\delta_{ij}(E^2 + H^2)\right]\Big/4\pi \tag{1-36}$$

where $\delta_{ij}$ denotes the usual Kronecker $\delta$-symbol. By comparison of equations (1-34) and (1-35) we see that the Maxwell stress tensor should be equal to the negative of the radiation pressure tensor; it is instructive to verify this conclusion by direct calculation.

Consider a plane wave propagating in direction $\mathbf{n}_0$; from the macroscopic definition of radiation pressure we have

$$\begin{aligned}\mathsf{P} &= c^{-1}\oint I\mathbf{nn}\, d\omega = c^{-1}I_0\oint \delta(\mathbf{n} - \mathbf{n}_0)\mathbf{nn}\, d\omega = c^{-1}I_0\mathbf{n}_0\mathbf{n}_0 \\ &= (E_0{}^2/8\pi)\mathbf{n}_0\mathbf{n}_0\end{aligned} \tag{1-37}$$

which should equal $\mathsf{T}^M$ for a plane wave. Choose the electromagnetic field to yield a Poynting vector $\mathbf{S}$ along $\mathbf{n}_0$; in addition to $(\theta_0, \phi_0)$ we must also specify the polarization of the wave via the angle $\psi_0$. Here $\psi_0$ measures the angle of rotation of $\mathbf{E}$ around $\mathbf{S}$ from the plane through $\mathbf{n}_0$ and $\mathbf{k}$ (the unit vector in the $z$-direction); see Figure 1-4. It is easy to see that

$$E_x = E_0(\sin\psi_0 \sin\phi_0 - \cos\psi_0 \cos\phi_0 \cos\theta_0) \tag{1-38a}$$

$$E_y = -E_0(\sin\psi_0 \cos\phi_0 + \cos\psi_0 \sin\phi_0 \cos\theta_0) \tag{1-38b}$$

$$E_z = E_0 \cos\psi_0 \sin\theta_0 \tag{1-38c}$$

*Exercise 1-7*: Derive expressions analogous to equation (1-38) for $(H_x, H_y, H_z)$.



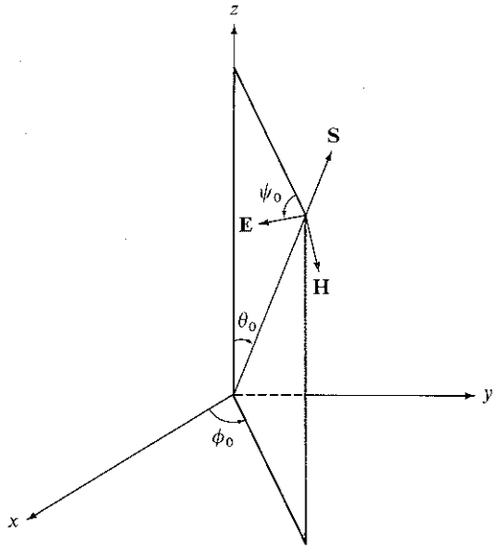



FIGURE 1-4
The plane electromagnetic wave generated by $\mathbf{E}$ and $\mathbf{H}$ is propagating along the Poynting vector $\mathbf{S}$ in direction $\mathbf{n}_0$. The angle $\psi_0$ measures the rotation of $\mathbf{E}$ around $\mathbf{S}$ out of the plane defined by $\mathbf{n}_0$ and $\mathbf{k}$, the unit vector in the $z$-direction.

Substitution of equations (1-38) and the corresponding equations for $\mathbf{H}$ into equation (1-36) yields components of $\mathsf{T}^M$; for example, for $T_{zz}{}^M$ we find

$$\begin{aligned}T_{zz}{}^M &= \left[E_z{}^2 + H_z{}^2 - \frac{1}{2}(E^2 + H^2)\right]\Big/4\pi = E_0{}^2(\sin^2\theta_0 - 1)/4\pi \\ &= -E_0{}^2\cos^2\theta_0/4\pi\end{aligned} \tag{1-39}$$

Averaging over time yields $\langle T_{zz}{}^M\rangle_T = -(E_0{}^2/8\pi)\cos^2\theta_0$ which is indeed $-P_{zz}$; note that the final result is independent of $\psi_0$.

*Exercise 1-8*: Calculate the remaining components of $\mathsf{T}^M$ and show that $\mathsf{T}^M = -\mathsf{P}$, independent of $\psi_0$.

The above results demonstrate that a complete correspondence exists between electromagnetic theory and the macroscopic or photon descriptions of the radiation field; we shall exploit this correspondence in a useful way in §§14-3 and 15-3 where we will be able to use the known Lorentz-transformation properties of electromagnetic field quantities to establish those of their radiation-field-description counterparts.

LIMITING CASES: SYMMETRY, ISOTROPY, EQUILIBRIUM, PLANE WAVES

In a one-dimensional planar or spherically-symmetric atmosphere the radiation field is azimuthally invariant, hence the pressure tensor becomes *diagonal*

$$\mathsf{P}(\mathbf{r}, \nu, t) = \begin{pmatrix} p_R & 0 & 0 \\ 0 & p_R & 0 \\ 0 & 0 & p_R \end{pmatrix} - \frac{1}{2}\begin{pmatrix} 3p_R - E_R & 0 & 0 \\ 0 & 3p_R - E_R & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{1-40}$$

where

$$p_R(z, \nu, t) \equiv (4\pi/c)K(z, \nu, t) \tag{1-41}$$

and, in turn,

$$K(z, \nu, t) \equiv \frac{1}{2}\int_{-1}^{1} I(z, \mu, \nu, t)\mu^2 \, d\mu \tag{1-42}$$

is the second moment of the radiation field in Eddington's notation.

*Exercise 1-9*: (a) Derive equation (1-40) for the conditions stated. (b) Show that the same expression for P is obtained in spherical symmetry relative to the orthogonal triad $(\hat{\boldsymbol{\phi}}, \hat{\boldsymbol{\theta}}, \hat{\mathbf{r}})$. (c) Verify that the tensor shown in equation (1-40) is consistent with equation (1-32).

It is clear from equation (1-40) that, for a *one-dimensional* atmosphere, only two scalars ($p_R$ and $E_R$) are sufficient to specify the full radiation pressure tensor. Further, for such atmospheres, derivatives with respect to $(x, y)$ or $(\theta, \phi)$, in the planar and spherical cases respectively, are identically zero, and the only nonvanishing components of the divergence of the radiation pressure tensor are

$$(\boldsymbol{\nabla} \cdot \mathsf{P})_z = \partial p_R(z, \nu, t)/\partial z \tag{1-43a}$$

in planar geometry—or, in spherical geometry,

$$(\boldsymbol{\nabla} \cdot \mathsf{P})_r = [\partial p_R(r, \nu, t)/\partial r] + [3p_R(r, \nu, t) - E_R(r, \nu, t)]/r \tag{1-43b}$$

In this book we shall confine attention strictly to one-dimensional problems, and with the exception of further formal development of the equations of radiation hydrodynamics in §15-3, the full tensor description of the radiation field will not be required; for ease of expression, we shall therefore refer to the single scalar $p_R$ as "the" radiation pressure.

*Exercise 1-10*: Show that, for any diagonal tensor A, in spherical coordinates $(\boldsymbol{\nabla} \cdot \mathsf{A})_r = (\partial A_{rr}/\partial r) + (2A_{rr} - A_{\theta\theta} - A_{\phi\phi})/r$; use this result to derive equation (1-43b).



In general, $p_R$ as defined by equation (1-41) will *not* equal $\bar{P}$ defined by equation (1-32), but the two become equal if the radiation field is isotropic. For an isotropic field, $I$ is independent of $\mu$, and equations (1-7) and (1-41) immediately yield $p_R = \frac{1}{3}E_R$, so that equation (1-40) reduces to

$$\mathsf{P}(\mathbf{r}, \nu, t) = \begin{pmatrix} p_R & 0 & 0 \\ 0 & p_R & 0 \\ 0 & 0 & p_R \end{pmatrix} \tag{1-44}$$

That is, *when the radiation field is isotropic*, the radiation pressure tensor is *diagonal and isotropic*, and may be replaced, for all purposes of computation, by a scalar *hydrostatic pressure*, or even eliminated entirely in terms of $E_R$. If in addition to being isotropic the radiation field has its *thermal equilibrium* value, then the monochromatic radiation pressure is

$$p_R^*(z, \nu, t) = \frac{1}{3}E_R^*(z, \nu, t) = (4\pi/3c)B_\nu(T) \tag{1-45}$$

and the total radiation pressure is

$$p_R^*(z, t) = \frac{1}{3}aT^4 \tag{1-46}$$

a result first obtained by thermodynamic arguments (**160**, 55; **565**, 123).

From equations (1-7) and (1-41) we see that $p_R$ is an average of $I(\mu)$ weighted by $\mu^2$ whereas $E_R$ is a straight average. If the radiation field becomes peaked in the direction of radiation flow out of the atmosphere, the larger values of $\mu$ become more heavily weighted in $p_R$ (recall $\mu = 1$ for $\theta = 0$) and $p_R$ will *exceed* its isotropic limit of $\frac{1}{3}E_R$. The most extreme departure from isotropy occurs when the radiation flows in a plane wave. For a wave in the outward direction we can write $I(z, \nu, \mu) = I(z, \nu)\,\delta(\mu - 1)$; then $K(z, \nu) = J(z, \nu) = H(z, \nu)$, and $p_R(z, \nu) = E_R(z, \nu)$. This extreme limit is approached in the outermost layers of very extended stellar envelopes (or in nebulae) in which the radiation field originates from a stellar surface that occupies only a very small solid angle as seen from the point in question.

*Exercise 1-11*: (a) Show that for a plane wave moving along one of the coordinate axes, the radiation pressure tensor has only one non-zero component. (b) Show that the pressure tensor is isotropic if the angular dependence of the radiation field is given by $I(\mu) = I_0 + I_1\mu$. This result is important because the radiation field is accurately described by an expression of the stated form in the *diffusion limit* which obtains at great depth in the atmosphere (cf. §2-5).

*Exercise 1-12*: Suppose an observer is at distance $r$ from the center of a star of radius $r_*$ which has a uniformly bright surface (i.e., $I$ is independent of $\mu$). Derive analytical expressions for $J$, $H$, and $K$ in terms of $\theta_* \equiv \sin^{-1}(r_*/r)$, and show that in the limit $(r/r_*) \to \infty$, $J = H = K$.

VARIABLE EDDINGTON FACTORS

From the results derived above it follows that the ratio $p_R(\mathbf{r}, \nu, t)/E_R(\mathbf{r}, \nu, t)$ or $K(\mathbf{r}, \nu, t)/J(\mathbf{r}, \nu, t)$ is a dimensionless number whose value is fixed by the degree of isotropy of the radiation field, and typically ranges between $\frac{1}{3}$ and 1. It will be shown later (§6-3) that this ratio can be used in certain numerical methods to reduce the number of independent variables in the transfer problem; further, it may be used to effect a closure of the system of moment equations derived from the transfer equation. It is useful, therefore, to define the *variable Eddington factor*

$$f(\mathbf{r}, \nu, t) \equiv K(\mathbf{r}, \nu, t)/J(\mathbf{r}, \nu, t) \tag{1-47}$$

or, in abbreviated notation $f_\nu = K_\nu/J_\nu$.

*Exercise 1-13*: (a) Consider an expansion of the form $I(\mu) = I_0 + \sum_n I_n \mu^n$; show that $f = \frac{1}{3}$ if the sum includes only *odd* powers $n$. (b) Suppose that $I(\mu) \equiv I_1$ for $(0 \leqslant \mu \leqslant 1)$ and $I(\mu) \equiv I_2$ for $(-1 \leqslant \mu \leqslant 0)$; show again that $f = \frac{1}{3}$. This representation of $I$ provides a rough description (the *two-stream approximation*) of a stellar radiation field, for we may let $I_2/I_1 \to 0$ at the surface and $I_2/I_1 \to 1$ at depth. (c) Show that, for a slab of infinite extent in $(x, y)$ and *finite* extent in $z$, $f$ may drop *below* $\frac{1}{3}$.

# 2

# *The Equation of Transfer*

As radiation passes through the gas composing a stellar atmosphere, it interacts with the material and is absorbed, emitted, and scattered repeatedly. These phenomena determine how *radiative transfer* occurs in the atmosphere. In this chapter, macroscopic quantities that define radiation–matter interactions are introduced (§2-1), and the *equation of transfer* (which describes the transport of radiation through the medium) is developed (§2-2). Using this equation, we can compute the emergent spectrum from a star, and calculate how the angle–frequency variation of the radiation field changes with depth in the atmosphere. The *time-dependent* equation of transfer will be derived in order to obtain moment equations (§2-3) that describe the *dynamical* behavior of the radiation field, but the discussion will then be restricted to *static* atmospheres in all subsequent work through Chapter 13. In Chapters 14 and 15 radiative transfer and its dynamical effects in *steady* (i.e., time-independent) *flows* will be considered.

## 2-1 The Interaction of Radiation with Matter

### DISTINCTION BETWEEN SCATTERING AND ABSORPTION–EMISSION PROCESSES

In the interaction of radiation with matter, energy may be removed from, or delivered into, the radiation field by a wide variety of physical processes. For the present, it is adequate to characterize these processes by macroscopic coefficients; as will be seen in Chapters 4 and 7, these coefficients are specified by atomic cross-sections and occupation numbers of energy levels of the constituents of the stellar material. It is worthwhile, from the outset, to make a distinction between "true" *absorption* and *emission* on one hand and the process of *scattering* on the other for, as we shall see repeatedly in the development of the theory, the physical nature of the interaction between the atmospheric material and radiation is quite different in these two cases. However, it is also important to realize that in spectral lines the dichotomy between these processes can be established uniquely only when we consider a transition between two specified atomic states, with no coupling to any other states allowed. As soon as *sequences* of transitions among several interacting states are considered, fundamental ambiguities arise, and it is no longer possible to describe a given line as an "absorption" or a "scattering" line in a rigorous way; nor would it be important or useful to do so. Nevertheless, it is fruitful to have at least an intuitive notion of the contrast between these two basic processes, obtained by consideration of some definite examples.

We may identify as *scattering* processes those in which a photon interacts with a scattering center (perhaps producing a change in the scatterer's internal excitation state) and emerges from the interaction in a new direction with (in general) a *slightly* altered energy. The essential point is that in this process, the energy of the photon is not converted into kinetic energy of particles in the gas. In contrast, we shall identify *absorption* processes as those in which the photon is destroyed by conversion of its energy (wholly or partly) into the thermal energy of the gas. In this process we say that the photon has been *thermalized*. The crucial physical point to note is that the local rate of energy emission in *scattering processes* depends mainly upon *the radiation field* (which may have originated at some other remotely situated point in the atmosphere) *and has only a weak connection with the local values of the thermodynamic properties* (e.g., temperature) *of the gas. Absorption processes*, on the other hand, *feed photon energy directly into the thermal kinetic energy of the gas, and hence are more intimately coupled to local thermodynamic properties of the material.* Conversely, the inverse of absorption, *thermal emission*, transfers energy *from the thermal pool* of the gas directly *into the radiation field.* Thermal absorption and emission processes thus tend to produce *local* equilibrium between the radiation and material; but scattering processes allow photons to move from one part of the atmosphere to some other part without coupling to local conditions, and thus tend to delocalize the control of the gas–radiation equilibration process, and to introduce global properties of the atmosphere (e.g., the presence of boundaries) into the problem.



To illustrate the ideas developed above, let us consider the following as typical examples of *scattering processes.*

(a) The interaction of a photon with an atom in bound state $a$ leading to the excitation of a higher bound state $b$ (the photon's energy being converted to internal excitation energy of the atom), followed by a direct return to state $a$ with the emission of a photon. In general the emitted photon will propagate in a different direction from that of the incident photon. Further, both the lower and upper states $a$ and $b$ of an atom in a radiating gas will not be perfectly sharp, but will have finite energy widths arising, for example, from the finite lifetime of each state produced by radiative decay, or from interactions of the atom with other particles of the plasma in which it is imbedded. Each of the bound states can, therefore, be considered to consist of a distribution of substates, with radiative transitions possible from any substate of one level to any substate of the other. Thus if the decay of the upper level occurs to a different substate of the lower level than that from which the excitation occurred, or if there is a redistribution of the excited electron from the original excited substate to some other substate (because, say, the atom suffers an elastic collision with another particle) then the emitted photon's energy may be slightly different from the incident photon's. Similarly, motions of the scattering centers with respect to the fixed laboratory frame can change the emitted photon's energy from the incident energy if the projection of the scatterer's velocity along the direction of propagation is different for the two photons, for then a differential Doppler shift can occur. (Example: imagine the incident photon to be moving in the *same* direction as $\mathbf{v}$ of the scatterer and the emitted photon to move in the *opposite* direction. The emitted photon will be *redshifted* by an amount $\Delta\nu = -2\nu_0 v/c$ relative to the incident photon). Changes in photon direction and frequency during scattering are described by *redistribution functions* (see below). Note that in this process no significant part of the photon energy is imparted to the material.

(b) Scattering of a photon by a free electron (Thomson or Compton scattering) or by an atom or molecule (Rayleigh scattering). Thomson scattering may be viewed as the result of the free charge oscillating in the electromagnetic field of the radiation, Compton scattering as a collision of a photon with a free charged particle, and Rayleigh scattering as a resonance of a permitted "oscillation" of the bound system with the field. The remarks made in (a) above concerning redistribution and lack of coupling of radiant energy to the thermal pool apply here as well.

Similarly, we may consider the following to be examples of *thermal absorption processes* (and their inverses as thermal emission).

(a) A photon is absorbed by an atom in a bound state, and ionizes the bound electron, allowing it to escape with finite kinetic energy into the continuum. In this process of *photoionization* or *bound–free absorption*, the photon is destroyed and the excess of its energy over the electron's binding energy goes initially into the electron's kinetic energy, and ultimately into the general thermal pool after the electron suffers elastic collisions that establish a thermal velocity distribution for the particles. The inverse process, of a free electron dropping to a bound state with the creation of a photon whose energy equals the sum of the electron's kinetic and binding energies, is called direct *radiative recombination.* These processes clearly transfer energy back and forth between the radiation field and the thermal pool of the material.

(b) A photon is absorbed by a free electron moving in the field of an ion, resulting in an alteration of the electron's kinetic energy relative to the ion. The electron then, classically speaking, moves off on a different (hyperbolic) orbit around the ion. This process is known as *free–free absorption* because the electron is unbound both before and after absorbing the photon. The inverse process, leading to the emission of a photon, is referred to as *bremsstrahlung.*

(c) A photon is absorbed by an atom, leading to a transition of an electron from one bound state to another; this process is called *photoexcitation* or *bound–bound absorption.* The atom is then de-excited by an inelastic collision with another particle. Energy is put into the motion of the atom and the collision partner and thereby ends up as part of the thermal pool. The photon is said to have been *destroyed* by a *collisional de-excitation.* The inverse process leads to the collisional creation of a photon at the expense of the thermal energy of the gas.

(d) Photoexcitation of an atom with subsequent *collisional ionization* of the excited atom into the continuum. Photon energy again contributes to the thermal energy of particles. The inverse processes is called (three-body) *collisional recombination.*

To illustrate the conceptual limitations of the kinds of arguments given above, let us now consider some ambiguous cases. Suppose an atom has three bound levels $a$, $b$, and $c$, in order of increasing energy, and a photo-excitation from $a$ to $c$ occurs. Then suppose that $c$ decays radiatively to $b$, and $b$ then decays radiatively to $a$; this process is called *fluorescence.* Here a single photon of energy $h\nu_{ac} = E_c - E_a$ is degraded into two photons of energies $h\nu_{ab} = E_b - E_a$ and $h\nu_{bc} = E_c - E_b$. Was the original photon scattered or absorbed? By our original definition it has not been "scattered" and, moreover, the new photons may have vastly different properties (e.g., probability of escape through the boundary surface) from the original, so



that the nonlocal behavior of the radiation field has been altered. On the other hand, no contribution has been made to the thermal energy of the gas. Alternatively, consider the same process, but now with a collisional de-excitation $c \rightarrow b$ followed by an emission $b \rightarrow a$. The original photon can be said to have been destroyed (absorbed); but is the emitted photon "thermally" emitted when most of the original energy was derived from the radiation field? Many other more complex and subtle cases may be constructed, which, taken together show the limits of usefulness of the absorption-vs-scattering description.

In fact, a truly consistent picture emerges only when we write down the full *equations of statistical equilibrium* (cf. Chapter 5), which describe all possible processes (both radiative and collisional) that couple an arbitrary state $i$ to some other state $j$, and solve these together with the equations (*transfer equations*) that describe how the radiation is absorbed, emitted, and transported through the atmosphere. To do this is, in general, quite difficult, and formulation of successful methods of solution of the problem will occupy the bulk of this book. (The full import of these comments will emerge only when the student has studied the material through Chapter 12; nevertheless they should be borne in mind at all stages of subsequent development).

### THE EXTINCTION COEFFICIENT

To describe the removal of energy from the radiation field by matter let us introduce a macroscopic coefficient $\chi(\mathbf{r}, \nu, t)$ called the *extinction coefficient*, or *opacity*, or sometimes (loosely) the *total absorption coefficient.* This coefficient is defined such that an element of material, of cross-section $dS$ and length $ds$, removes from a beam with specific intensity $I(\mathbf{r}, \mathbf{n}, \nu, t)$, incident normal to $dS$ and propagating into a solid angle $d\omega$, an amount of energy

$$\delta E = \chi(\mathbf{r}, \mathbf{n}, \nu, t) I(\mathbf{r}, \mathbf{n}, \nu, t)\, dS\, ds\, d\omega\, d\nu\, dt \tag{2-1}$$

within a frequency band $d\nu$ in a time $dt$. The extinction coefficient is the product of an atomic absorption cross-section ($\text{cm}^2$) and the number density of absorbers ($\text{cm}^{-3}$) summed over all states that can interact with photons of frequency $\nu$. The dimensions of $\chi$ are $\text{cm}^{-1}$, and $(1/\chi)$ gives a measure of the distance over which a photon can propagate before it is removed from the beam—i.e., a *photon mean-free-path* (cf. §2-2).

The frequency variation of $\chi$ may be extremely complicated, and may include thousands or millions of transitions (bound–bound, bound–free, and free–free). For *static media* in which there are no preferred directions imposed on an atomic scale (e.g., by a magnetic field), the *opacity is isotropic.* For *moving media*, the *opacity has an angular dependence* introduced by the

Doppler shift that radiation experiences in the fluid frame relative to its original frequency in the stationary laboratory frame; this Doppler shift obviously depends on the projection of the velocity vector onto the direction of the incident beam. In what follows we consider only static atmospheres.

As outlined earlier in this section, it is sometimes useful to distinguish between "absorption" and "scattering"; hence we introduce volume coefficients $\kappa(\mathbf{r}, \nu, t)$ and $\sigma(\mathbf{r}, \nu, t)$ that describe [via equation (2-1)] the rate at which energy is removed from the beam by "true absorption" and "scattering," respectively. The total extinction is given by

$$\chi(\mathbf{r}, \nu, t) = \kappa(\mathbf{r}, \nu, t) + \sigma(\mathbf{r}, \nu, t) \tag{2-2}$$

That is, both processes are assumed to occur independently and to add linearly. In actual practice $\chi$ is sufficient to describe energy removed from the beam; the distinction between $\kappa$ and $\sigma$ is useful mainly in defining the emission coefficient.

In the calculation of $\chi$ it is necessary to include a correction for *stimulated emission* (see §§4-1 and 4-3). This is a quantum process in which radiation *induces* a downward transition from the upper state at a rate proportional to the product of a cross-section, the upper-state population, and the specific intensity. Because the process is proportional to $I(\mathbf{r}, \mathbf{n}, \nu, t)$ and effectively cancels out some of the opacity, it is convenient to include it in the definition of $\chi$. Stimulated emission occurs only when the emitting system exists in a *definite* upper state (whether bound or free). There is thus *no* stimulated emission in Thomson scattering (free electrons) or Rayleigh scattering (involves *virtual* states) but there *is* stimulated emission in spectrum lines, even if they are described with a "scattering" coefficient.

If we know the value of $\chi(\mathbf{r}, \nu, t)$ [or of $\kappa(\mathbf{r}, \nu, t)$ and $\sigma(\mathbf{r}, \nu, t)$], we have a complete *macroscopic* description of the rate at which material removes energy from a beam of radiation. But it is crucial to emphasize that the "completeness" of the description is *illusory*. The reason this unpleasant comment must be made is that the simple picture we obtain from equation (2-1) glosses over the fact that the level populations, which "determine" the rate of energy removal from the radiation field by their contribution to $\chi$, are, in turn, *determined by* the radiation field via photoexcitations, photoionizations, radiative emission, radiative recombination, and related processes. Thus, in reality, the interaction of the radiation field with the absorbing material is *nonlinear*. The problem just described still remains (though more subtly) even if it is assumed that we can calculate level populations by local application of thermodynamic equilibrium relations that depend only on the density and temperature. (This is the so-called *local thermodynamic equilibrium* or *LTE* approximation.) The reason is that the temperature is determined by overall balance between energy emitted and energy absorbed



by the material, and thus by the nature of the radiation field and its response to the global properties of the atmosphere (e.g., boundaries, scattering, gradients, etc.). The remarks made in this paragraph carry over with equal force in the macroscopic description of emission. Again, the full significance of these remarks will emerge only with considerable further development (see particularly §§5-1 and 5-3).

### THE EMISSION COEFFICIENT

To describe the emission of radiation from the stellar material, we introduce a macroscopic *emission coefficient* or *emissivity* $\eta(\mathbf{r}, \mathbf{n}, \nu, t)$ defined such that the amount of energy released from an element of material of cross-section $dS$ and length $ds$, into a solid angle $d\omega$, within a frequency band $d\nu$, in direction $\mathbf{n}$ in a time interval $dt$, is

$$\delta E = \eta(\mathbf{r}, \mathbf{n}, \nu, t)\, dS\, ds\, d\omega\, d\nu\, dt \tag{2-3}$$

The dimensions of $\eta$ are ergs $\text{cm}^{-3}$ $\text{sr}^{-1}$ $\text{hz}^{-1}$ $\text{sec}^{-1}$. As was true for the opacity, *thermal* emissivity is *isotropic* for *static* media (without imposed preferred directions) but is *angle-dependent* for *moving* material owing to Doppler-shift effects. For radiation emitted in scattering processes, there is normally an explicit angle-dependence, even for static media. The emissivity is calculated by summing products of upper-state populations and transition probabilities over all relevant processes that can release a photon at frequency $\nu$. In writing the transfer equation we shall usually use the unembellished symbol $\eta$ to denote the total emissitivity; if electron scattering terms appear explicitly in the same equation, $\eta$ will then denote all other emission. Subscripts "$c$" and "$l$" may on occasion be used to denote continua and lines, respectively. Again, we must realize that the simplicity of this description is deceptive, for the reasons given above in the discussion of the extinction coefficient.

An important relation exists between the emission and absorption coefficients in the case of strict *thermodynamic equilibrium* (T.E.). If we consider an adiabatic enclosure in steady-state equilibrium containing a homogeneous medium, we know that the material will have the same temperature $T$ throughout (otherwise it would be possible to devise processes to extract work from the temperature gradient, in violation of the second law of thermodynamics). Further, we may expect the radiation field to be isotropic and homogeneous throughout the enclosure (including at the surface of the walls), for if it were not, beams traveling in opposite directions would not be exactly similar and a directional transport of energy would result, from which work could be extracted, again in violation of the second law of thermodynamics. Consideration of the energy absorbed and emitted in

angle–frequency ranges $d\omega\, d\nu$ by an element of material, in time $dt$, now shows that if a steady-state thermal equilibrium is to be achieved (no *net* gain or loss of energy by the matter), the thermal emission must be given by

$$\eta^t(\nu) = \kappa(\nu) I(\mathbf{n}, \nu) \tag{2-4}$$

which is known as *Kirchhoff's law*. For an enclosure in strict T.E. at temperature $T$ the intensity of the radiation field is given by the Planck function $B_\nu(T)$, so that

$$\eta^*(\nu) = \kappa^*(\nu) B_\nu(T) \tag{2-5}$$

which is the *Kirchhoff–Planck relation*. This result has been obtained without reference to the composition of the material and is valid (in T.E.) for *all* materials. [See the excellent discussion of the interaction of matter and radiation in T.E. in (**160**, 199–206) and in an article by Milne in (**416**, 93–96).]

Strictly speaking, the Kirchhoff–Planck law applies only in the case of a system in T.E. But if the material is subject only to *small* gradients over the mean free path a photon can travel before it is destroyed and thermalized by a collisional process (as is true, e.g., in the interior of a star), then we could expect equation (2-5) to be valid to a high degree of approximation at *local* values of the thermodynamic variables specifying the state of the material. In such a case we write

$$\eta^t(\mathbf{r}, \nu, t) = \kappa^*(\mathbf{r}, \nu, t) B_\nu[T(\mathbf{r}, t)] \tag{2-6}$$

The hypothesis of *local thermodynamic equilibrium* (or LTE) just mentioned makes the *assumption* that the occupation numbers of bound and free states of the material, the opacity, the emissivity—indeed *all* of the thermodynamic properties of the material—are the same as their T.E. values at the local values of $T$ and density, throughout the *entire* atmosphere, out to the outermost regions. Only the radiation field is allowed to depart from its T.E. value of $B_\nu[T(\mathbf{r})]$, and is obtained from a solution of the transfer equation. Such an approach is manifestly *internally inconsistent*, although LTE expressions remain valid for certain quantities even in the general case. For example, equation (2-6) is a valid expression for the *continuum* emission coefficient even in the presence of departures from LTE so long as the velocity distribution of recombining (or, for free–free emission, colliding) particles is Maxwellian; the equation is *not* valid for *line* emission, and further, the LTE formula for the opacity is not correct. The use of LTE is a computational expedient that simplifies the calculation of models of stellar atmospheres, and has been widely applied. (We shall employ it in places to provide a prototype with which we can introduce basic mathematical techniques of solving transfer problems, and will discuss models built assuming LTE in §§7-2 through 7-4.) But it must be stressed that stellar atmospheres are regions



in which there are *large* gradients of material properties, and an open boundary through which radiation freely escapes; the radiation field is therefore highly anisotropic and has a markedly non-Planckian character. One might argue for using LTE, even though the radiation field is obtained from solving a *nonlocal* transfer equation, if one could show that some mechanism, specifically collisions among the particles, *enforced* LTE populations. As already mentioned above, and as we shall see in detail in §5-3 and in Chapters 7 and 11–14, this will not, in general, be the case. Rather, the radiation field determines the state of the material, and hence equation (2-6) becomes invalid; in the end we must carry through a general analysis in which we specify the thermodynamic state of the gas and the distribution function of the radiation field simultaneously by solving the coupled equations of transfer and statistical equilibrium.

Let us now consider radiation *scattered* by the material. For simplicity of notation, we suppress explicit reference to $t$, though all of the quantities may be time-dependent. As described earlier, in scattering processes both the direction and frequency of a photon may change. These changes are described by a *redistribution function*

$$R(\nu', \mathbf{n}'; \nu, \mathbf{n})\, d\nu'\, d\nu (d\omega'/4\pi)(d\omega/4\pi)$$

which gives the *joint probability* that a photon will be scattered *from* direction $\mathbf{n}'$ in solid angle $d\omega'$ and frequency range $(\nu', \nu' + d\nu')$ *into* solid angle $d\omega$ in direction $\mathbf{n}$ and frequency range $(\nu, \nu + d\nu)$. We shall derive redistribution functions and discuss them in detail in Chapter 13, but it is helpful to mention some of the general properties of these functions at the present time. We shall *normalize* $R$ such that

$$(4\pi)^{-2} \oint d\omega' \oint d\omega \int_0^\infty d\nu' \int_0^\infty d\nu\, R(\nu', \mathbf{n}'; \nu, \mathbf{n}) = 1 \tag{2-7}$$

The redistribution function contains within it both a normalized *absorption profile* $\phi(\nu)$, and a normalized *emission profile* $\psi(\nu)$ for the scattering process. From the physical definition given above, it is evident that if we integrate over all emitted frequencies and angles we must obtain the probability for absorption from solid angle $d\omega'$ and frequency range $\nu'$—i.e., $\phi(\nu')\, d\nu'\, d\omega'/4\pi$. Thus

$$\phi(\nu') = (4\pi)^{-1} \oint d\omega \int_0^\infty d\nu\, R(\nu', \mathbf{n}'; \nu, \mathbf{n}) \tag{2-8}$$

which by virtue of equation (2-7) is normalized such that $\int \phi(\nu')\, d\nu' = 1$. If $\sigma_0(\mathbf{r})$ denotes the total scattering coefficient, we may write $\sigma(\mathbf{r}, \nu') = \sigma_0(\mathbf{r})\phi(\nu')$. The joint probability that an amount of energy $\sigma_0(\mathbf{r}) I(\mathbf{r}, \mathbf{n}', \nu')$ will be removed from the beam in solid angle $d\omega'$ at frequency $\nu'$, and scattered into $d\omega$ at frequency $\nu$, is $\sigma_0(\mathbf{r}) R(\nu', \mathbf{n}'; \nu, \mathbf{n}) I(\mathbf{r}, \mathbf{n}', \nu')\, d\nu'\, d\nu (d\omega'/4\pi)(d\omega/4\pi)$. Hence if we

integrate over all incoming frequencies and angles, we find the total amount of energy emitted at frequency $\nu$ into solid angle $d\omega$, namely

$$\eta^S(\mathbf{r}, \mathbf{n}, \nu)\, d\nu(d\omega/4\pi) = \sigma_0(\mathbf{r})\, d\nu(d\omega/4\pi) \oint (d\omega'/4\pi) \int_0^\infty d\nu'\, R(\nu', \mathbf{n}'; \nu, \mathbf{n}) I(\mathbf{r}, \mathbf{n}', \nu') \tag{2-9}$$

*Exercise 2-1*: Show that the *total* emission rate $\int d\nu \oint d\omega\, \eta(\mathbf{r}, \mathbf{n}, \nu)$ equals the *total* energy removed from the beam $4\pi\sigma_0(\mathbf{r}) \int \phi(\nu') J(\mathbf{r}, \nu')\, d\nu'$—i.e., that the scattering process is *conservative*.

Equation (2-9) gives the full angle–frequency dependence of the emission profile. It is usually difficult to treat radiative transfer problems in the degree of generality implied here, and useful simplifications of the problem can be made. For example, if we are primarily interested in redistribution in frequency and not in angle, we could assume that $I(\mathbf{r}, \mathbf{n}, \nu)$ is nearly isotropic, and replace it in equation (2-9) with $J(\mathbf{r}, \nu)$. Then the emission into $d\nu\, d\omega$ is

$$\eta^S(\mathbf{r}, \nu) = \sigma_0(\mathbf{r}) \int_0^\infty R(\nu', \nu) J(\mathbf{r}, \nu')\, d\nu' \tag{2-10}$$

where the *angle-averaged redistribution function*

$$R(\nu', \nu) \equiv (4\pi)^{-1} \oint R(\nu', \mathbf{n}'; \nu, \mathbf{n})\, d\omega' = (4\pi)^{-1} \oint R(\nu', \mathbf{n}'; \nu, \mathbf{n})\, d\omega \tag{2-11}$$

gives the redistribution probability from $(\nu', \nu' + d\nu')$ to $(\nu, \nu + d\nu)$ and is normalized such that

$$\int_0^\infty d\nu' \int_0^\infty d\nu\, R(\nu', \nu) = \int_0^\infty \phi(\nu')\, d\nu' = 1 \tag{2-12}$$

The function $R(\nu', \nu)$ is rendered independent of angle as a result of integrating over either $d\omega'$ *or* $d\omega$; this follows from the fact that (cf §13-2) $R(\nu', \mathbf{n}'; \nu, \mathbf{n})$ depends only on the angle between $\mathbf{n}'$ and $\mathbf{n}$. Equation (2-10) provides an extremely useful approximation in line transfer problems because the crucial phenomenon there is the frequency diffusion of photons from the opaque line core (where they are trapped) to the more transparent line wings (whence they may escape from the atmosphere at depths where $I$ is, in fact, very nearly isotropic). In the angle-averaged approximation the emission profile

$$\psi(\nu) \equiv \eta(\mathbf{r}, \nu) \Big/ \int_0^\infty \eta(\mathbf{r}, \nu)\, d\nu$$

is given by

$$\psi(\nu) = \int_0^\infty R(\nu', \nu) J(\mathbf{r}, \nu')\, d\nu' \Big/ \int_0^\infty \phi(\nu') J(\mathbf{r}, \nu')\, d\nu' \tag{2-13}$$



which shows that the distribution of emitted photons depends upon the frequency profile of the incoming radiation.

In the limiting case that the intensity is independent of frequency, we obtain *natural excitation*, with

$$\psi^*(\nu) = \int_0^\infty R(\nu', \nu)\, d\nu' \tag{2-14}$$

If we have $R(\nu', \nu) = R(\nu, \nu')$, as occurs in most cases of interest (cf. §13-3), then $\psi^*(\nu) = \phi(\nu)$; that is, *for natural excitation, the emission and absorption profiles are identical* (a result that is *not* true in general!). Natural excitation prevails, of course, in T.E., and it is usually thought of in that context. There are, however, other physical circumstances in which the result $\psi(\nu) = \phi(\nu)$ is recovered. In particular, suppose that there is a complete reshuffling of atoms in their excited state in such a way that there is *no correlation between the frequencies of incoming and scattered photons*; then both have frequencies *independently* distributed over the absorption profile $\phi(\nu)$. This situation is referred to as *complete redistribution*, or *complete noncoherence*. A good approximation to this case occurs, for instance, when atoms are so strongly perturbed by collisions during the scattering process that the excited electrons are randomly redistributed over substates of the upper level. In this case, both the absorption and emission probabilities *independently* are proportional to the number of substates available at each frequency within the line [i.e., to $\phi(\nu)$ itself], and the joint absorption–emission probability $R(\nu', \nu)$ is the product of these two independent distributions—i.e., $R(\nu', \nu) = \phi(\nu')\phi(\nu)$. For *complete redistribution*, the emissivity is

$$\eta^S(\mathbf{r}, \nu) = \sigma_0(\mathbf{r})\phi(\nu) \int_0^\infty \phi(\nu') J(\mathbf{r}, \nu')\, d\nu' \tag{2-15}$$

from which we see clearly that the emission and absorption profiles are identical. Complete redistribution is also a good approximation within the Doppler core of a spectrum line, and actually provides an excellent first approximation in line transfer problems. We shall, in fact, assume complete redistribution in our discussion of line formation until Chapter 13.

Another class of problems arises when we focus attention on the angular redistribution of the emitted radiation, but assume that the scattering is essentially *coherent* (i.e., $\nu' = \nu$). This is the situation of interest, e.g., in scattering of light by large particles in a planetary (including earth's) atmosphere. We can then write

$$R(\nu', \mathbf{n}'; \nu, \mathbf{n}) = g(\mathbf{n}', \mathbf{n})\phi(\nu')\, \delta(\nu - \nu') \tag{2-16}$$

where $\delta$ is the Dirac function and $g$ is an *angular phase function* normalized such that

$$(4\pi)^{-1} \oint g(\mathbf{n}', \mathbf{n})\, d\omega' = 1 \tag{2-17}$$

Two important phase functions are those for isotropic scattering

$$g(\mathbf{n}', \mathbf{n}) \equiv 1 \tag{2-18}$$

and the *dipole* phase function (which applies for Thomson and Rayleigh scattering)

$$g(\mathbf{n}', \mathbf{n}) = \frac{3}{4}(1 + \cos^2 \Phi) \tag{2-19}$$

where $\cos \Phi = \mathbf{n}' \cdot \mathbf{n}$. The phase functions for scattering by large particles (i.e., whose size is comparable to a wavelength of light) are often extremely complicated and show large, rapid variations as a function of angle (**312**; **359**, Chap. 4).

For *coherent scattering*, equation (2-9) reduces to

$$\eta^S(\mathbf{r}, \mathbf{n}, \nu) = \sigma(\mathbf{r}, \nu) \oint I(\mathbf{r}, \mathbf{n}', \nu) g(\mathbf{n}', \mathbf{n}) \frac{d\omega'}{4\pi} \tag{2-20}$$

In a spectrum line, coherent scattering would occur only if the lower state of the line were sharp, the upper state were not perturbed during emission, and the scattering atoms were at rest in the observer's frame. This is not the case, however, and line scattering is much more accurately described by complete redistribution (except in the far line wings where $\phi$ varies slowly over the range corresponding to a Doppler shift). On the other hand, for continuum scattering (e.g., by electrons) the frequency distribution of radiation is smooth and essentially constant over the typical frequency shifts occurring in the scattering process. For this reason continuum scattering processes are customarily treated as if they were coherent (though this may be inadequate near a spectrum line). Moreover, as the angular redistribution effects from a dipole phase function are usually very small in a stellar atmosphere, it is customary to assume that continuum scattering is also isotropic, and to write

$$\eta^S(\mathbf{r}, \nu) = \sigma(\mathbf{r}, \nu) J(\mathbf{r}, \nu) \tag{2-21}$$

## 2-2 The Transfer Equation

### DERIVATION

Let us now consider the problem of radiative transport. Choose an inertial coordinate system and examine the flow of energy through a fixed volume element in a definite time interval. Let us assume that the radiation field is, in general, time-dependent. If we suppose the material to be at rest, then both $\chi$ and $\eta$ will be isotropic (unless we consider anisotropic scattering). In moving material one must account for changes in photon frequency and



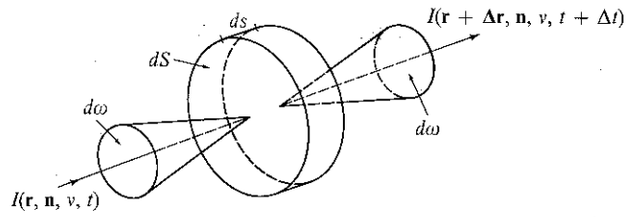



FIGURE 2-1
Element of absorbing and emitting material considered in derivation of transfer equation.

direction (Doppler shift and aberration) resulting from the transformation between the laboratory frame and the fluid frame. These effects depend upon $\mathbf{n} \cdot \mathbf{v}$; hence both $\chi$ and $\eta$ will have an explicit angle-dependence in this case. Now calculate the energy in a frequency interval $d\nu$, passing in a time $dt$ through a volume element of length $ds$ and cross-section $dS$ oriented normal to a ray traveling in direction $\mathbf{n}$ into solid angle $d\omega$ (see Figure 2-1). The difference between the amount of energy that emerges (at position $\mathbf{r} + \Delta\mathbf{r}$ at time $t + \Delta t$) and that incident (at $\mathbf{r}$ and $\mathbf{t}$) must equal the amount created by emission from the material in the volume minus the amount absorbed. That is,

$$\begin{aligned}[I(\mathbf{r} + \Delta\mathbf{r}, \mathbf{n}, \nu, t + \Delta t) - I(\mathbf{r}, \mathbf{n}, \nu, t)]\, dS\, d\omega\, d\nu\, dt \\ = [\eta(\mathbf{r}, \mathbf{n}, \nu, t) - \chi(\mathbf{r}, \mathbf{n}, \nu, t) I(\mathbf{r}, \mathbf{n}, \nu, t)]\, ds\, dS\, d\omega\, d\nu\, dt\end{aligned} \tag{2-22}$$

Let $s$ denote the path-length along the ray; then $\Delta t = \Delta s/c$, and

$$I(\mathbf{r} + \Delta\mathbf{r}, \mathbf{n}, \nu, t + \Delta t) = I(\mathbf{r}, \mathbf{n}, \nu, t) + [c^{-1}(\partial I/\partial t) + (\partial I/\partial s)]\, ds \tag{2-23}$$

Substituting equation (2-23) into equation (2-22) we have the *transfer equation*

$$[c^{-1}(\partial/\partial t) + (\partial/\partial s)] I(\mathbf{r}, \mathbf{n}, \nu, t) = \eta(\mathbf{r}, \mathbf{n}, \nu, t) - \chi(\mathbf{r}, \mathbf{n}, \nu, t) I(\mathbf{r}, \mathbf{n}, \nu, t) \tag{2-24}$$

The derivative along the ray may be re-expressed in terms of an orthogonal coordinate system:

$$\frac{\partial I}{\partial s} = \left(\frac{\partial x}{\partial s}\right)\left(\frac{\partial I}{\partial x}\right) + \left(\frac{\partial y}{\partial s}\right)\left(\frac{\partial I}{\partial y}\right) + \left(\frac{\partial z}{\partial s}\right)\left(\frac{\partial I}{\partial z}\right) = n_x \frac{\partial I}{\partial x} + n_y \frac{\partial I}{\partial y} + n_z \frac{\partial I}{\partial z} \tag{2-25}$$

where $(n_x, n_y, n_z)$ are the components of the unit vector $\mathbf{n}$. We may thus rewrite equation (2-24) as

$$[c^{-1}(\partial/\partial t) + (\mathbf{n} \cdot \nabla)] I(\mathbf{r}, \mathbf{n}, \nu, t) = \eta(\mathbf{r}, \mathbf{n}, \nu, t) - \chi(\mathbf{r}, \mathbf{n}, \nu, t) I(\mathbf{r}, \mathbf{n}, \nu, t) \tag{2-26}$$

For a *one-dimensional planar atmosphere* $n_z = (dz/ds) = \cos\theta = \mu$; further, the derivatives $(\partial/\partial x)$ and $(\partial/\partial y)$ are identically zero, and we obtain

$$[c^{-1}(\partial/\partial t) + \mu(\partial/\partial z)]I(z, \mathbf{n}, \nu, t) = \eta(z, \mathbf{n}, \nu, t) - \chi(z, \mathbf{n}, \nu, t)I(z, \mathbf{n}, \nu, t) \tag{2-27}$$

or, for the *time-independent* case,

$$\mu[\partial I(z, \mathbf{n}, \nu)/\partial z] = \eta(z, \mathbf{n}, \nu) - \chi(z, \mathbf{n}, \nu)I(z, \mathbf{n}, \nu) \tag{2-28}$$

Equation (2-28) is the "standard" transfer equation for plane-parallel model-atmospheres calculations; the coordinate $z$ increases *upward* in the atmosphere (i.e., toward an external observer). For *static* media, the specification $(z, \mathbf{n}, \nu)$ in $\eta$ and $\chi$ may be reduced to $(z, \nu)$ only. Note that if $\eta$ and $\chi$ are *given*, equation (2-28) is an *ordinary differential equation*, which may be solved for all relevant choices of $\mu$ and $\nu$. When $\eta$ includes *scattering* terms, the transfer equation becomes an *integro-differential equation* containing angle and frequency integrals of $I$.

### THE TRANSFER EQUATION AS A BOLTZMANN EQUATION

The basic equation describing particle transport in kinetic theory is the *Boltzmann equation*; we shall now show that the transfer equation is just the Boltzmann equation for photons. Suppose we have a particle distribution function $f(\mathbf{r}, \mathbf{p}, t)$ that gives the number density of particles in the phase volume element $(\mathbf{r}, \mathbf{r} + \mathbf{dr})$, $(\mathbf{p}, \mathbf{p} + \mathbf{dp})$. We follow the evolution of $f$ within a particular phase-space element for a time interval $dt$, in which $\mathbf{r} \to \mathbf{r} + \mathbf{v}\,dt$ and $\mathbf{p} \to \mathbf{p} + \mathbf{F}\,dt$, where $\mathbf{F}$ denotes externally imposed forces acting on the particles. The phase-space element evolves from

$$(d^3r)_0(d^3p)_0 \to (d^3r)(d^3p) = J[(d^3r)_0(d^3p)_0]$$

where $J$ is the *Jacobian* of the transformation.

*Exercise 2-2*: Show that to first order in $dt$ the Jacobian of the transformation of a phase volume element is $J = 1$.

In view of the result of Exercise 2-2 we see that the phase-space element is *deformed*, but its phase volume is *unchanged*. If all external forces $\mathbf{F}$ are continuous, then the deformation of the phase-space element is continuous, and all particles originally within the volume remain there; as the volume itself is unchanged, the particle density is unchanged. But if, in addition, collisions occur, individual particles may be reshuffled from one element of phase space to another "discontinuously"; i.e., their neighbors may be totally unaffected during the same time interval. Therefore, the change in the particle



number density within a phase-space element must equal the *net* number introduced into the element by collisions; i.e.,

$$\frac{\partial f}{\partial t} + \left(\frac{\partial x}{\partial t}\right)\left(\frac{\partial f}{\partial x}\right) + \left(\frac{\partial y}{\partial t}\right)\left(\frac{\partial f}{\partial y}\right) + \left(\frac{\partial z}{\partial t}\right)\left(\frac{\partial f}{\partial z}\right) + F_x\left(\frac{\partial f}{\partial p_x}\right) + F_y\left(\frac{\partial f}{\partial p_y}\right) + F_z\left(\frac{\partial f}{\partial p_z}\right) = \left(\frac{Df}{Dt}\right)_{\text{coll}} \tag{2-29}$$

or, in more compact notation

$$(\partial f/\partial t) + (\mathbf{v}\cdot\nabla)f + (\mathbf{F}\cdot\nabla_p)f = (Df/Dt)_{\text{coll}} \tag{2-30}$$

For a "gas" consisting of photons (with rest-mass zero), in the absence of general relativistic effects, $\mathbf{F} \equiv 0$, and photon propagation in an inertial frame occurs in straight lines with $\mathbf{v} = c\mathbf{n}$, while the frequency remains constant. The distribution function $f_R$ can be written in terms of the specific intensity by means of equation (1-2). The analogues of "collisions" are photon interactions with the material, and the net number of photons introduced into the volume will be the energy emitted minus the energy absorbed, divided by the energy per photon. Thus for photons equation (2-30) becomes

$$(chv)^{-1}[(\partial I/\partial t) + c(\mathbf{n}\cdot\nabla)I] = (\eta - \chi I)/(h\nu) \tag{2-31}$$

which is identical to the transfer equation (2-26). In effect the transfer equation is a Boltzmann equation for a fluid that is subject to no external forces but which suffers strong collisional effects. As will be seen in §2-3, the *moments* of the transfer equation yield *dynamical equations* for the radiation field, just as moments of the Boltzmann equation for a gas lead to equations of hydrodynamics.

### SPHERICAL GEOMETRY

In a *spherically symmetric* medium, the specific intensity will be independent of the coordinates $\Theta$ and $\Phi$ of the triplet $(r, \Theta, \Phi)$ which specifies a position in the atmosphere, and of the azimuthal angle $\phi$ of the pair $(\theta, \phi)$ which specifies the direction of the beam relative to the local outward normal $\hat{\mathbf{r}}$. Thus $I(\mathbf{r}, \mathbf{n}, \nu, t)$ reduces to $I(r, \theta, \nu, t)$. In writing the transfer equation, starting from the general form of equation (2-24), we must now account for the variation of $\theta$ along a displacement, and employ the general form $\mathbf{ds} = dr\,\hat{\mathbf{r}} + r\,d\theta\,\hat{\boldsymbol{\theta}}$. As is clear from the geometry of the situation (see Figure 2-2), $dr = \cos\theta\,ds$, while $r\,d\theta = -\sin\theta\,ds$ (note that $d\theta \leqslant 0$ for *any* $ds$), so that

$$(\partial/\partial s) \to \cos\theta(\partial/\partial r) - r^{-1}\sin\theta(\partial/\partial\theta) = \mu(\partial/\partial r) + r^{-1}(1 - \mu^2)(\partial/\partial\mu) \tag{2-32}$$

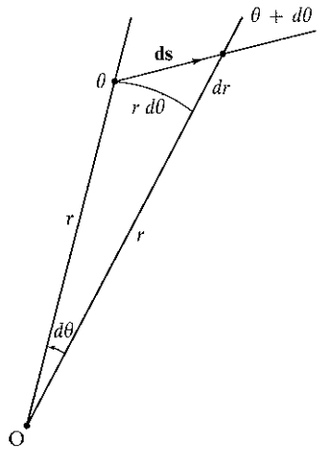



FIGURE 2-2
Geometric relations among variables used in derivation of transfer equation in a spherically symmetric medium.

where, as usual, $\mu = \cos\theta$. Hence the transfer equation for a *spherically symmetric* atmosphere is

$$[c^{-1}(\partial/\partial t) + \mu(\partial/\partial r) + r^{-1}(1 - \mu^2)(\partial/\partial\mu)]I(r, \mu, \nu, t) = \eta(r, \nu, t) - \chi(r, \nu, t)I(r, \mu, \nu, t) \quad (2\text{-}33)$$

which simplifies in an obvious way in the time-independent case. Note that now, even with $\eta$ and $\chi$ specified, equation (2-33) is a *partial differential equation* in $r$ and $\mu$. However, this added complexity can be avoided by using the (straight-line) characteristic paths that reduce the spatial operator to a single derivative with respect to pathlength (see §7-6). In fact, equation (2-33) is not structurally different from equation (2-28) and can be solved almost as easily.

*Exercise 2-3*: (a) Consider an atmosphere which is *axially symmetric* but not spherically symmetric (e.g., a rotationally flattened star). Show that now $I(\mathbf{r}, \mathbf{n}) = I(r, \Theta, \theta, \phi)$. (b) For the general case where $I(\mathbf{r}, \mathbf{n}) = I(r, \Theta, \Phi, \theta, \phi)$ (e.g., a rotationally flattened star illuminated by a companion), show that the transfer equation in spherical coordinates, accounting for all the spatial derivatives, is

$$c^{-1}(\partial I/\partial t) + \mu(\partial I/\partial r) + (\gamma/r)(\partial I/\partial\Theta) + (\sigma/r\sin\Theta)(\partial I/\partial\Phi) + r^{-1}(1 - \mu^2)(\partial I/\partial\mu) - (\sigma\cot\Theta/r)(\partial I/\partial\phi) = \eta - \chi I$$

where $\gamma \equiv \cos\phi\sin\theta$ and $\sigma \equiv \sin\phi\sin\theta$.

### OPTICAL DEPTH AND THE SOURCE FUNCTION

For the remainder of §2-2 let us confine attention to the time-independent planar transfer equation (2-28). Writing $d\tau(z, \nu) \equiv -\chi(z, \nu)\,dz$, we define an



*optical depth scale* $\tau(z, \nu)$ which gives the integrated absorptivity of the material along the line of sight as

$$\tau(z, \nu) = \int_z^{z_{\max}} \chi(z', \nu)\,dz' \quad (2\text{-}34)$$

The negative sign is introduced so that the optical depth increases *inward* into the atmosphere from zero at the surface (where $z = z_{\max}$), and thus provides a measure of how deeply an outside observer can see into the material [cf. equations (2-47) and (2-52)]. Recalling that $\chi^{-1}$ is the photon mean-free-path, we recognize that $\tau(z, \nu)$ is just the *number of photon mean-free-paths* at frequency $\nu$ along the line of sight from $z_{\max}$ to $z$. In addition, we define the *source function* to be the ratio of the total emissivity to total opacity,

$$S(z, \nu) = \eta(z, \nu)/\chi(z, \nu) \quad (2\text{-}35)$$

To simplify the notation we shall for the present suppress explicit reference to $z$ and $\mu$, and denote frequency dependence with a subscript $\nu$. The equation of transfer may then be written in its standard form

$$\mu(\partial I_\nu/\partial\tau_\nu) = I_\nu - S_\nu \quad (2\text{-}36)$$

From the discussion of §2-1 we can write prototype expressions for the source function which we use to study methods of solving equation (2-36). Suppose first we have strict LTE. Then from equation (2-6) we have

$$S_\nu = B_\nu \quad (2\text{-}37)$$

If we have a contribution from thermal absorption and emission plus a contribution from a coherent, isotropic, continuum scattering term (say from Thomson scattering by free electrons or from Rayleigh scattering) then we could write

$$\chi_\nu = \kappa_\nu + \sigma_\nu \quad (2\text{-}38)$$

and

$$S_\nu = (\kappa_\nu B_\nu + \sigma_\nu J_\nu)/(\kappa_\nu + \sigma_\nu) \quad (2\text{-}39)$$

For a spectrum line with an overlapping background continuum we have

$$\chi_\nu = \chi_c + \chi_l(\nu) = \chi_c + \chi_l\phi_\nu \quad (2\text{-}40)$$

where $\chi_c$ and $\chi_l$ denote the continuum and line opacities, respectively. If we assume that a fraction $\varepsilon$ of the line emission occurs from thermal processes and the remainder is given by angle-averaged complete redistribution [equation (2-15)], we can write

$$\eta_\nu = \chi_c B_\nu + \chi_l\phi_\nu\left[\varepsilon B_\nu + (1 - \varepsilon)\int\phi_\nu J_\nu\,d\nu\right] \quad (2\text{-}41)$$

and

$$S_\nu = \left(\frac{r + \varepsilon\phi_\nu}{r + \phi_\nu}\right) B_\nu + \left[\frac{(1 - \varepsilon)\phi_\nu}{r + \phi_\nu}\right] \int \phi_\nu J_\nu \, d\nu = \xi_\nu B_\nu + (1 - \xi_\nu) \int \phi_\nu J_\nu \, d\nu \tag{2-42}$$

where $r \equiv \chi_c/\chi_l$. (We may ignore the frequency variation of the continuum over a line-width.) In equations (2-39) and (2-42) we have examples of the explicit appearance of integrals of the radiation field over angle and frequency, which shows the integro-differential nature of the transfer equation. It must be stressed that the source functions in equations (2-37), (2-39), and (2-42) are *only illustrative*; they are based on essentially heuristic arguments, and a physically rigorous formulation can be provided only after the equations of statistical equilibrium (Chapter 5) are developed.

### BOUNDARY CONDITIONS

Solution of the transfer equation requires the specification of *boundary conditions*. Two problems of fundamental astrophysical importance are (a) the *finite slab* (in planar geometry) or *shell* (spherical geometry), and (b) a medium (e.g., a stellar atmosphere) that has an open boundary on one side but is so optically thick that it can be imagined to extend to infinity on the other side—the *semi-infinite* atmosphere.

For the finite slab of total geometrical thickness $Z$ and total optical depth $T_\nu$ (defined to be zero on the side nearest the observer), a unique solution is obtained if the intensity incident on both faces of the slab is specified. Writing $\theta$ for the angle between a ray and the normal directed toward the observer, and $\mu \equiv \cos\theta$, we must specify the two functions $I^+$ and $I^-$ such that

$$I(\tau_\nu = 0, \mu, \nu) = I^-(\mu, \nu), \quad (-1 \leqslant \mu \leqslant 0) \tag{2-43}$$

at the upper boundary, and

$$I(\tau_\nu = T_\nu, \mu, \nu) = I^+(\mu, \nu), \quad (0 \leqslant \mu \leqslant 1) \tag{2-44}$$

at the lower boundary. For a shell of outer radius $R$ and inner radius $r_c$, equation (2-43) still applies at $r = R$.

*Exercise 2-4*: (a) At the inner boundary of a spherical shell, $r = r_c$, show that the lower boundary condition is given by $I(r_c, +\mu, \nu) \equiv I(r_c, -\mu, \nu)$ if the central volume $r \leqslant r_c$ is *void*; this is Milne's "planetary nebula" boundary condition. (b) If the volume contains a *point source* (a star) of intensity $I_0$, show that the result in (a) must be augmented by $I(r_c, +1, \nu) = I_0(\nu)\,\delta(\mu - 1)$. (c) Extend results (a) and (b) to the case of the volume being partially filled by an *opaque source* on the range $0 \leqslant r \leqslant r_*$, $(r_* \leqslant r_c)$ with intensity $I_0(\mu', \nu)$ where $\mu'$ is the angle cosine *at the opaque surface*; this case simulates an envelope around a star. (d) Show in case (a) above that the flux is *identically zero* at all points $r \leqslant r_c$.



In the semi-infinite case (planar or spherical), the radiation field incident upon the upper boundary must be specified by an equation of the form of (2-43); for stellar atmospheres work, it is customary to assume $I^- \equiv 0$ (clearly this would *not* be done, e.g., in a binary system). For a lower boundary condition one may replace equation (2-44) by a *boundedness condition* in analytical work where the limit $\tau_\nu \to \infty$ is taken. Specifically we impose the requirement that

$$\lim_{\tau \to \infty} I(\tau_\nu, \mu, \nu) e^{-\tau_\nu/\mu} = 0 \tag{2-45}$$

The reasons for this particular choice will become clear in the discussion below. Alternatively, at great depth in the atmosphere we may write $I(\tau_\nu, \mu, \nu)$ in terms of the *local* value of $S_\nu$ and its gradient, or may specify the *flux*; these conditions follow naturally from physical considerations in the *diffusion limit* where the photon mean free path is much smaller than its optical depth from the surface (see §2-5).

### SIMPLE EXAMPLES

Before writing the formal solution of the equation of transfer, it is instructive to consider a few simple examples in planar geometry.

(a) Suppose no material is present. Then $\chi_\nu = \eta_\nu = 0$, and equation (2-28) reduces to $(\partial I_\nu/\partial z) = 0$ or $I_\nu =$ constant. This result is consistent with the proof in §1-1 of the invariance of the specific intensity when no sources or sinks are present.

(b) Suppose that the material emits at frequency $\nu$, but cannot absorb. Then equation (2-28) is $\mu(\partial I_\nu/\partial z) = \eta_\nu$, and for a finite slab the emergent intensity is given by

$$I(Z, \mu, \nu) = \mu^{-1} \int_0^Z \eta(z, \nu)\, dz + I^+(0, \mu, \nu) \tag{2-46}$$

The physical situation described above occurs in the formation of optically forbidden lines in nebulae. Atoms may be excited to *metastable levels* by collisions; because nebular densities are so low, the chances of a second collision leading to de-excitation are very small, so the atoms can remain unperturbed in these levels for long periods of time, and large numbers of atoms may accumulate in these states. Eventually some of the atoms decay via "forbidden" transitions, which have very small but nonzero transition probabilities, and emit photons. Because the line is forbidden, the probability of reabsorption is negligible, and the photon escapes. Thus photons are created

at the expense of the energy in the thermal pool, and none are destroyed by absorption.

(c) Suppose that radiation is absorbed but not emitted by the material. Then $\mu(\partial I_\nu/\partial z) = -\chi_\nu I_\nu$, and defining $d\tau_\nu = -\chi_\nu\, dz$, the emergent intensity from a finite slab of total optical thickness $T_\nu$ is

$$I(\tau_\nu = 0, \mu, \nu) = I^+(T_\nu, \mu, \nu)\exp(-T_\nu/\mu) \tag{2-47}$$

Equation (2-47) applies, for example, to radiation passing through a filter in which photons are absorbed and degraded into photons of some very different frequency (e.g., extreme far-infrared heat radiation) before being re-emitted, or are destroyed and converted into kinetic energy of the particles in the absorbing medium.

### FORMAL SOLUTION

Let us now obtain a formal solution of the equation of transfer; we confine attention exclusively to planar geometry. Regarding $S$ as given, equation (2-36) is a linear first-order differential equation with constant coefficients, and must therefore have an integrating factor. The integrating factor is easily shown to be $\exp(-\tau_\nu/\mu)$, so that

$$\partial[I_\nu \exp(-\tau_\nu/\mu)]/\partial\tau_\nu = -\mu^{-1}S_\nu \exp(-\tau_\nu/\mu) \tag{2-48}$$

Integration of equation (2-48) yields

$$I(\tau_1, \mu, \nu) = I(\tau_2, \mu, \nu)e^{-(\tau_2-\tau_1)/\mu} + \mu^{-1}\int_{\tau_1}^{\tau_2} S_\nu(t)e^{-(t-\tau_1)/\mu}\, dt \tag{2-49}$$

If $S_\nu$ is given, equation (2-49) provides a complete solution of the planar transfer problem.

We may apply equation (2-49) at an *arbitrary interior point in a semi-infinite atmosphere*. For $\mu \geqslant 0$ (outgoing radiation), set $\tau_1 = \tau_\nu$ and $\tau_2 = \infty$, and impose the boundedness condition of equation (2-45); the result is

$$I(\tau_\nu, \mu, \nu) = \int_{\tau_\nu}^{\infty} S_\nu(t)e^{-(t-\tau_\nu)/\mu}\, dt/\mu, \quad (0 \leqslant \mu \leqslant 1) \tag{2-50}$$

For incoming radiation, $(-1 \leqslant \mu \leqslant 0)$, set $\tau_2 = 0$ and apply the upper boundary condition $I^- \equiv 0$; we then obtain

$$I(\tau_\nu, \mu, \nu) = \int_0^{\tau_\nu} S_\nu(t)e^{-(\tau_\nu - t)/(-\mu)}\, dt/(-\mu), \quad (-1 \leqslant \mu \leqslant 0) \tag{2-51}$$

One of the most important applications of equation (2-50) is the expression for the *emergent intensity* seen by an external observer ($\tau = 0$),

$$I(0, \mu, \nu) = \int_0^{\infty} S_\nu(t)e^{-t/\mu}\, dt/\mu \tag{2-52}$$



The physical significance of equation (2-52) is that the emergent intensity is a weighted average of the source function along the line of sight; the weighting function is the fraction of the energy emitted at each depth that penetrates to the surface along a ray whose optical slant-length is $(\tau/\mu)$. Mathematically speaking, equation (2-52) shows that the specific intensity is the *Laplace transform* of the source function, a property that may be used to solve for $S$ in certain classes of problems.

Considerable insight can be gained by supposing the source function is a linear function of depth: $S_\nu(\tau_\nu) = S_{0\nu} + S_{1\nu}\tau_\nu$; then equation (2-52) yields

$$I(0, \mu, \nu) = S_{0\nu} + S_{1\nu}\mu = S_\nu(\tau_\nu = \mu) \tag{2-53}$$

which is known as the *Eddington–Barbier relation*. This result states that *the emergent intensity is characteristic of the value of the source function at about optical depth unity along the line of sight* (recall that $\tau_\nu$ represents the *normal* optical depth, so that *slant* optical depth unity for a ray penetrating at angle $\cos^{-1}\mu$ occurs at $\tau_\nu = \mu$). The Eddington–Barbier relation has been widely applied in empirical analyses of the solar and stellar spectra, and provides a basic conceptual framework for many interpretive methods. In the case of the sun, when we observe, at a fixed frequency, the variation in intensity from disk center ($\mu = 1$) to the limb ($\mu \to 0$) we can (in principle) infer information about $S_\nu(\tau_\nu)$ for $0 \lesssim \tau_\nu \lesssim 1$. For stars, we cannot observe the center-to-limb variation, but it is clear that if we observe at different frequencies (e.g., within a spectral line), we encounter unit optical depth in higher layers for frequencies with high opacity (e.g., line-center) and in deeper layers for frequencies with low opacity (e.g., line-wings). If we then know something about the frequency-variation of $S_\nu$, we can infer information about its depth-variation; for example, in LTE, $S_\nu = B_\nu$ and the frequency-variation over a narrow line can be ignored, so that one can, in principle, infer the run of the temperature with depth. Although it is extremely useful conceptually, the Eddington–Barbier relation should *not* be applied indiscriminately and used literally to argue that $I(0, \mu, \nu)$ is identical to $S_\nu(\tau_\nu = \mu)$, because (a) there are always significant contributions to $I_\nu(0)$ from *other* depths (i.e., there is an intrinsic "fuzziness" in the problem), and (b) the assumptions from which it follows may not be valid. A detailed critique of the limitations of the Eddington–Barbier relation can be found in (**18**, 121–130) and (**20**, 20–30).

> *Exercise 2-5*: Suppose the source function is to be represented by a power-series expansion about the point $\tau_*$; i.e., $S(\tau) \approx S(\tau_*) + S'(\tau_*)(\tau - \tau_*) + \frac{1}{2}S''(\tau_*)(\tau - \tau_*)^2$. Calculate the emergent intensity and show that the choice $\tau_* = \mu$ is "optimum" in the sense that it *eliminates* the contribution of $S'$ and *minimizes* the contribution of $S''$ to $I(0, \mu)$.

Another instructive example to consider is a finite slab of optical thickness $T$, within which $S$ is constant, and upon which there is no incident radiation.

The normally emergent intensity is $I(0, 1) = S(1 - e^{-T})$. For $T \gg 1, I = S$. This result is sensible physically, for the radiation that emerges consists of those photons emitted over a mean-free-path from the surface; the emission rate is $\eta$ and the mean-free-path is $\chi^{-1}$, hence $I = \eta/\chi = S$. For $T \ll 1$, $e^{-T} \approx 1 - T$, so $I \approx ST$. Again this result is physically reasonable, for in the optically thin case we can see through the entire volume; hence the energy emitted (per unit area) must be the emissivity per unit volume $\eta$ times the total path-length $Z$, or $I = \eta Z = (\eta/\chi)(\chi Z) = ST$. Note that this result is consistent with equation (2-46).

THE SCHWARZSCHILD–MILNE EQUATIONS

By integration of the formal solution for the specific intensity over angle, concise expressions for the moments of the radiation field may be derived. Thus by substitution of equations (2-50) and (2-51) into equation (1-4) we have for the mean intensity

$$J_\nu(\tau_\nu) \equiv \frac{1}{2}\int_{-1}^{1} I_\nu(\tau_\nu, \mu)\, d\mu = \frac{1}{2}\left[\int_0^1 d\mu \int_{\tau_\nu}^{\infty} S_\nu(t) e^{-(t-\tau_\nu)/\mu} \frac{dt}{\mu} + \int_{-1}^{0} d\mu \int_0^{\tau_\nu} S_\nu(t) e^{-(\tau_\nu - t)/(-\mu)} \frac{dt}{(-\mu)}\right] \tag{2-54}$$

Equation (2-54) is reduced to a more useful form by interchanging the order of integration, and making the substitution $w = \pm 1/\mu$ in the first and second integrals respectively. Then

$$J_\nu(\tau_\nu) = \frac{1}{2}\left[\int_{\tau_\nu}^{\infty} dt\, S_\nu(t) \int_1^\infty dw \frac{e^{-w(t-\tau_\nu)}}{w} + \int_0^{\tau_\nu} dt\, S_\nu(t) \int_1^\infty dw \frac{e^{-w(\tau_\nu - t)}}{w}\right] \tag{2-55}$$

The integrals over $w$ are of a standard form and are called the first *exponential integral*. In general, for integer-values of $n$, the $n$th exponential integral is defined as

$$E_n(x) \equiv \int_1^\infty t^{-n} e^{-xt}\, dt = x^{n-1} \int_x^\infty t^{-n} e^{-t}\, dt \tag{2-56}$$

In terms of $E_1(x)$, equation (2-55) can be rewritten as

$$J_\nu(\tau_\nu) = \frac{1}{2}\int_0^\infty S_\nu(t_\nu) E_1|t_\nu - \tau_\nu| dt_\nu \tag{2-57}$$

Equation (2-57) was first derived by K. Schwarzschild and is named in his honor; Schwarzschild's paper (**416**, 35) is one of the foundation stones of radiative transfer theory, and merits careful reading. Because the integral appearing in equation (2-57) occurs so often in radiative transfer theory it has been abbreviated to an operator notation:



$$\Lambda_\tau[f(t)] \equiv \frac{1}{2}\int_0^\infty f(t) E_1|t - \tau| dt \tag{2-58}$$

*Exercise 2-6*: (a) Show that equations (2-50) and (2-51) are equivalent to

$$I(\mathbf{r}, \mathbf{n}) = \int_0^{s_{\max}} \eta(\mathbf{r}') \exp[-\tau(\mathbf{r}, \mathbf{r}')]\, d|\mathbf{r}' - \mathbf{r}|$$

where $\mathbf{r}'(s) \equiv \mathbf{r} - s\mathbf{n}$, $\tau(\mathbf{r}', \mathbf{r}) \equiv \int_0^{|\mathbf{r}'-\mathbf{r}|} \chi[\mathbf{r}'(s)]\, ds$, and $s_{\max}$ is the distance along the ray to any boundary surface in the direction $(-\mathbf{n})$; $s_{\max} = \infty$ for outward-directed rays in a semi-infinite medium. (b) Substitute the above result into the definition of $J(\mathbf{r})$ [equation (1-3)] to derive *Peierls's equation*:

$$J(\mathbf{r}) = (4\pi)^{-1} \int_V \{\eta(\mathbf{r}') \exp[-\tau(\mathbf{r}', \mathbf{r})]/|\mathbf{r}' - \mathbf{r}|^2\}\, d^3r'$$

where $V$ denotes the entire volume containing material.

By an analysis similar to that used to obtain equation (2-57) we can derive expressions for $F_\nu$ and $K_\nu$ first obtained by Milne (**416**, 77):

$$F_\nu(\tau_\nu) = 2\int_{\tau_\nu}^\infty S_\nu(t_\nu) E_2(t_\nu - \tau_\nu)\, dt_\nu - 2\int_0^{\tau_\nu} S_\nu(t_\nu) E_2(\tau_\nu - t_\nu)\, dt_\nu \tag{2-59}$$

and

$$K_\nu(\tau_\nu) = \frac{1}{2}\int_0^\infty S_\nu(t_\nu) E_3|t_\nu - \tau_\nu|\, dt_\nu \tag{2-60}$$

We also define the corresponding operators

$$\Phi_\tau[f(t)] \equiv 2\int_\tau^\infty f(t) E_2(t - \tau)\, dt - 2\int_0^\tau f(t) E_2(\tau - t)\, dt \tag{2-61}$$

and

$$X_\tau[f(t)] \equiv 2\int_0^\infty f(t) E_3|t - \tau|\, dt \tag{2-62}$$

*Exercise 2-7*: Derive equations (2-59) and (2-60).

The mathematical properties of the exponential integrals are discussed in detail in (**4**, 228–231) and (**161**, Appendix I), and the properties of the $\Lambda$, $\Phi$, and $X$ operators are discussed in (**361**, Chap. 2). A few of the most important results are mentioned in the following exercise.

*Exercise 2-8*: (a) Differentiate equation (2-56) to prove $E_n'(x) = -E_{n-1}(x)$. (b) Integrate equation (2-56) by parts to show that $E_n(x) = [e^{-x} - xE_{n-1}(x)]/(n-1)$ for $n > 1$. (c) Show that the asymptotic behavior $(x \gg 1)$ of $E_n(x)$ is $E_n(x) \sim e^{-x}/x$.

Some interesting physical insight can be gained by considering a linear source function $S(\tau) = a + b\tau$; with the results of Exercise 2-8, it is easy to show that

$$\Lambda_\tau(a + bt) = (a + b\tau) + \frac{1}{2}[bE_3(\tau) - aE_2(\tau)] \tag{2-63}$$

$$\Phi_\tau(a + bt) = \frac{4}{3}b + 2[aE_3(\tau) - bE_4(\tau)] \tag{2-64}$$

and

$$X_\tau(a + bt) = \frac{4}{3}(a + b\tau) + 2[bE_5(\tau) - aE_4(\tau)] \tag{2-65}$$

*Exercise 2-9*: Verify equations (2-63) through (2-65).

From the expressions written above (which will be applied in Chapter 3 and later work), we can note some important results. First, because the exponential integrals decay asymptotically as $e^{-x}/x$, it is evident from equation (2-63) that $J(\tau) = \Lambda_\tau(S)$ strongly approaches the local value of $S(\tau)$ for $\tau \gg 1$; that is, the $\Lambda$-operator tends to reproduce the source function at depth. In contrast, at the surface, $E_2(0) = 1$ and $E_3(0) = \frac{1}{2}$, and we obtain $\Lambda_0(S) = \frac{1}{2}a + \frac{1}{4}b$; in particular, if $b \equiv 0$, then $J(0) = \frac{1}{2}a = \frac{1}{2}S$. Physically this reflects the fact that $J$ at the surface is the average over a hemisphere containing no radiation (none incident from empty space) and a hemisphere in which $I = S$ (no gradient in $S$). When a gradient is present, $J(0)$ may lie above or below $S(0)$ depending on the sign and magnitude of the gradient. In a general way, $J$ will depart most strongly from $S$ at the surface. Second, we see from equation (2-64) that for $\tau \gg 1$, $H \to \frac{1}{3}b$. That is, the flux at depth depends only on the gradient of the source function (see also §2-5). The flux at the surface is $H(0) = \frac{1}{4}a + \frac{1}{6}b$; clearly the surface flux will be larger, the faster the source function rises inward. Note also that the effect of a gradient, relative to the case $b = 0$, is larger for $H(0)$ than for $J(0)$. (Why?)

Finally, it must be stressed that the solution of the transfer equation as given by equations (2-50) and (2-51) or (2-57) and (2-59) is only formal, and its apparent simplicity is illusory. For example, suppose the source function contains a scattering term as in equation (2-39) or (2-42). Then it is clear that the source function, which is required to compute the radiation field, *depends upon* the radiation field, and hence upon the solution of the transfer equation. In such cases we cannot calculate $I$ or $J$ by simple quadrature, but must find the solution of an *integral equation* or of the corresponding differential equation.



*Exercise 2-10*: Assume a source-function of the form of equation (2-39). Show that $J_\nu(\tau_\nu)$ satisfies the *integral equation*

$$J_\nu(\tau_\nu) = \Lambda_{\tau_\nu}[\rho_\nu(t_\nu)J_\nu(t_\nu)] + \Lambda_{\tau_\nu}\{[1 - \rho_\nu(t_\nu)]B_\nu(t_\nu)\}$$

where $\rho_\nu \equiv \sigma_\nu/(\kappa_\nu + \sigma_\nu)$

Complications of this kind are also introduced by other physical constraints on the solution. For example, we shall see in §2-4 that the requirement of energy balance couples together the radiation field and the temperature structure of the atmosphere. Thus even if strict LTE is assumed, and $S_\nu$ is set equal to $B_\nu$ and presumed to be independent of $J_\nu$, we are not in general able to specify the run of $T(\tau_\nu)$ and hence of $B_\nu[T(\tau_\nu)]$ until the radiation field is already known. Here again the source function required to compute the radiation field depends upon the field itself. When the more general non-LTE situation is considered, these problems become more subtle and complex, for now material properties (opacity, emissivity, etc.) depend directly upon the radiation field, for it determines occupation-numbers in atomic states. The transfer problem, in short, is fundamentally nonlinear in real stellar atmospheres. The development of methods to treat complicated interactions of the types just described will occupy a major part of this book.

## 2-3 Moments of the Transfer Equation

The angular moments of the transfer equation yield results of deep physical significance and great mathematical utility. The basic time-dependent transfer equation (2-26) may be rewritten as

$$c^{-1}[\partial I(\mathbf{r}, \mathbf{n}, \nu, t)/\partial t] + \sum_j n_j[\partial I(\mathbf{r}, \mathbf{n}, \nu, t)/\partial x_j] = \eta(\mathbf{r}, \mathbf{n}, \nu, t) - \chi(\mathbf{r}, \mathbf{n}, \nu, t)I(\mathbf{r}, \mathbf{n}, \nu, t) \tag{2-66}$$

where $x_j$ denotes the $j$th cartesian coordinate. To obtain the *zero-order moment equation*, we integrate equation (2-66) over all solid angles $d\omega$, and use the definitions in equations (1-4) and (1-19) to write

$$(4\pi/c)[\partial J(\mathbf{r}, \nu, t)/\partial t] + \nabla \cdot \mathscr{F}(\mathbf{r}, \nu, t) = \oint [\eta(\mathbf{r}, \mathbf{n}, \nu, t) - \chi(\mathbf{r}, \mathbf{n}, \nu, t)I(\mathbf{r}, \mathbf{n}, \nu, t)]\, d\omega \tag{2-67}$$

where we have allowed for the possibility that $\chi$ and $\eta$ may depend upon angle (as they will for moving media). If we integrate equation (2-67) over

all frequencies, and recall equation (1-8), we find

$$[\partial E_R(\mathbf{r}, t)/\partial t] + \nabla \cdot \mathscr{F}(\mathbf{r}, t) = \int_0^\infty d\nu \oint d\omega[\eta(\mathbf{r}, \mathbf{n}, \nu, t) - \chi(\mathbf{r}, \mathbf{n}, \nu, t)I(\mathbf{r}, \mathbf{n}, \nu, t)] \quad (2\text{-}68)$$

This is an *energy equation* for the radiation field, and it closely resembles the standard energy equation for a moving fluid (cf. §15-1). The terms in the equation have a straightforward physical interpretation: the time rate of change of the radiation energy density equals (a) the total energy emitted into the field by the material, (b) minus the total absorbed from the field by the material, (c) minus the net flow of energy through the surface containing a volume element (the divergence of the flux). If the medium is *static*, $\chi$ and $\eta$ are *isotropic* and the integrals on the righthand side simplify to yield

$$[\partial E_R(\mathbf{r}, t)/\partial t] + \nabla \cdot \mathscr{F}(\mathbf{r}, t) = 4\pi \int_0^\infty [\eta(\mathbf{r}, \nu, t) - \chi(\mathbf{r}, \nu, t)J(\mathbf{r}, \nu, t)]\, d\nu \quad (2\text{-}69)$$

For a *time-independent* radiation field in a *one-dimensional planar static medium*, equation (2-67) reduces to the "standard" result

$$[\partial H(z, \nu)/\partial z] = \eta(z, \nu) - \chi(z, \nu)J(z, \nu) \quad (2\text{-}70)$$

or, abbreviating the notation and using equations (2-34) and (2-35),

$$(\partial H_\nu/\partial \tau_\nu) = J_\nu - S_\nu \quad (2\text{-}71)$$

For *spherical geometry* (in a time-independent static atmosphere), by use of the appropriate expression for the divergence, we find

$$r^{-2}[\partial(r^2 H_\nu)/\partial r] = \eta_\nu - \chi_\nu J_\nu \quad (2\text{-}72)$$

Equations (2-70) through (2-72) will be employed repeatedly to obtain solutions of the transfer equation, and equation (2-69) will be used to develop the equations of radiation hydrodynamics (cf. §15-3).

*Exercise 2-11:* Derive equation (2-72) directly from equation (2-33).

The *first-order-moment equation* with respect to the $i$th coordinate axis is obtained by multiplying equation (2-66) by $n_i$ and integrating over $(d\omega/c)$, which yields

$$c^{-2}[\partial \mathscr{F}_i(\mathbf{r}, \nu, t)/\partial t] + \sum_j [\partial P_{ij}(\mathbf{r}, \nu, t)/\partial x_j] = c^{-1} \oint [\eta(\mathbf{r}, \mathbf{n}, \nu, t) - \chi(\mathbf{r}, \mathbf{n}, \nu, t)I(\mathbf{r}, \mathbf{n}, \nu, t)]n_i\, d\omega \quad (2\text{-}73)$$



or, in vector notation,

$$c^{-2}[\partial \mathscr{F}(\mathbf{r}, \nu, t)/\partial t] + \nabla \cdot \mathsf{P}(\mathbf{r}, \nu, t) = c^{-1} \oint [\eta(\mathbf{r}, \mathbf{n}, \nu, t) - \chi(\mathbf{r}, \mathbf{n}, \nu, t)I(\mathbf{r}, \mathbf{n}, \nu, t)]\mathbf{n}\, d\omega \quad (2\text{-}74)$$

Recalling that the momentum density in the radiation field is $\mathbf{G}_R = c^{-2}\mathscr{F}$ (see §1-3), we see that equation (2-74) is analogous to the hydrodynamical equations of motion, and may be viewed as a *dynamical equation* for the momentum in the radiation field at frequency $\nu$. When we integrate over all frequencies, we obtain a *total momentum equation* for the radiation, which we shall use in the equations of radiation hydrodynamics (cf. Chapter 15):

$$c^{-2}[\partial \mathscr{F}(\mathbf{r}, t)/\partial t] + \nabla \cdot \mathsf{P}(\mathbf{r}, t) = c^{-1} \int_0^\infty d\nu \oint d\omega[\eta(\mathbf{r}, \mathbf{n}, \nu, t) - \chi(\mathbf{r}, \mathbf{n}, \nu, t)I(\mathbf{r}, \mathbf{n}, \nu, t)]\mathbf{n} \quad (2\text{-}75)$$

Equation (2-75) states that the time rate of change of the total momentum density in the radiation field is equal to the negative of the volume force exerted by radiation stresses (cf. §1-4) plus a term that must represent the net momentum gain (or loss) from interacting with the material [see further discussion following equation (2-76)]. As in equation (2-68), equation (2-75) allows for the possibility of material motions. If the medium is at rest, then the integral over $\eta$ vanishes (which merely states that the net momentum loss from the material by *isotropic* emission is zero, as is physically obvious), while the second term reduces to an integral over the flux:

$$c^{-2}[\partial \mathscr{F}(\mathbf{r}, t)/\partial t] + \nabla \cdot \mathsf{P}(\mathbf{r}, t) = -c^{-1} \int_0^\infty \chi(\mathbf{r}, \nu, t)\mathscr{F}(\mathbf{r}, \nu, t)\, d\nu \quad (2\text{-}76)$$

The physical significance of the integral on the righthand side of equation (2-76) can readily be seen by the following argument. Consider a beam of specific intensity $I$, entering an element of absorbing material of surface area $dS$, at an angle $\theta$ relative to the normal. The energy absorbed by material of opacity $\chi$ from a solid angle $d\omega$ and frequency band $d\nu$ in time $dt$ is

$$dE = \chi I\, dS \cos\theta\, ds\, d\omega\, d\nu\, dt$$

where $ds = (dz/\cos\theta)$ is the slant-length of the ray through the element of thickness $dz$ along the normal. The component of momentum deposited in the material along the direction of the normal is $c^{-1}\, dE \cos\theta$; hence the momentum deposition per unit volume per unit time is

$$(c^{-1} \cos\theta\, dE)/(dz\, dS\, dt) = c^{-1}\chi I \cos\theta\, d\omega\, d\nu$$

If we sum over all angles and frequencies, we obtain precisely the integral in equation (2-76). We have thus shown that the integral is the *radiation force, per unit volume, on the material*; this interpretation is clearly compatible with the overall physical meaning of the equation.

For a *time-independent* radiation field in a *one-dimensional static planar medium*, equation (2-74) reduces to

$$[\partial p_R(z, \nu)/\partial z] = -(4\pi/c)\chi(z, \nu)H(z, \nu) \tag{2-77a}$$

or, integrating over all frequencies

$$[\partial p_R(z)/\partial z] = -(4\pi/c)\int_0^\infty \chi(z, \nu)H(z, \nu)\, d\nu \tag{2-77b}$$

Alternatively we can write

$$[\partial K(z, \nu)/\partial z] = -\chi(z, \nu)H(z, \nu) \tag{2-78}$$

or

$$(\partial K_\nu/\partial \tau_\nu) = H_\nu \tag{2-79}$$

In *spherical geometry*, under the same assumptions, equation (2-74) reduces, with the aid of the expression derived in Exercise 1-10 for $(\nabla \cdot \mathsf{P})_r$, to

$$(\partial K_\nu/\partial r) + r^{-1}(3K_\nu - J_\nu) = -\chi_\nu H_\nu \tag{2-80}$$

*Exercise 2-12*: Derive equation (2-80) directly from equation (2-33).

Thus far we have examined the moment equations primarily from the standpoint of their dynamical significance; but in the time-independent case they may also be used as tools to solve the transfer equation. By the introduction of moments, the angle-variable is eliminated and the dimensionality of the system to be solved is reduced. As we have seen in §2-2, the mean intensity can be determined from the solution of an integral equation (see Exercise 2-10). This gives the source function, from which the higher moments (e.g., the flux) can be determined by quadrature. The question now arises whether we can solve the moment equations as differential equations. Examination of equations (2-71) and (2-79) immediately reveals an essential difficulty: the moment equation of order $n$ always involves the moment of order $n + 1$, hence there is always one more variable than there are equations to determine them! This difficulty is known as the *closure problem*: one *additional* relation among the moments must somehow be obtained to "close" the system. For solving transfer equations, a variety of methods exist that employ moments of arbitrarily high order, and introduce *ad hoc* closure relations [see, e.g., (**361**, 90–101; **365**)]. However, in this book attention will be confined entirely to the moments $J_\nu$, $H_\nu$, and $K_\nu$ (an exception appears in §14-3), and the system will be closed (see §6-3) by eliminating $K_\nu$ in terms of $J_\nu$ and



the variable Eddington factor $f_\nu$—i.e., $K_\nu = f_\nu J_\nu$. The factor $f_\nu$ is obtained by *iteration* and allows us to effect an *approximate closure* of the *exact system* (if the iteration *converges*, the closure is also exact). Alternatively, it will be shown in §6-3 that the transfer equation can be rewritten in terms of *angle-dependent* mean-intensity-like and flux-like variables, and that *exact closure* of an angle-dependent equation that resembles a moment equation can be obtained. This equation is easily discretized and solved. In sum, the solution of the transfer equation in terms of moments or equivalent variables *can* be effected by differential-equation techniques of great generality and power.

## 2-4 The Condition of Radiative Equilibrium

Deep within the interior of a star, nuclear reactions release a flux of energy that diffuses outward, passes ultimately through the star's atmosphere, and emerges as observable radiation. In normal stars there is no creation of energy within the atmosphere itself; the atmosphere merely *transports* outward the total energy it receives. In a time-independent transport process, the frequency distribution of the radiation, or the partitioning of energy between radiative and nonradiative modes of transfer, may be altered; but *the energy flux as a whole is rigorously conserved.*

There are two basic modes of energy transport in those atmospheric layers in which spectrum-formation takes place: radiative and convective (or some other hydrodynamical mode). In these layers conduction is ineffective and can be ignored (it becomes important in *coronae* at temperatures of the order of $10^6$ °K). When all of the energy is transported by radiation, we have what is called *radiative equilibrium*; conversely, pure convective transport is called *convective equilibrium.* Whether or not radiative transport prevails over convection is determined by the *stability* of the atmosphere against convective motions.

The criterion for the stability of radiative transport was first enunciated by K. Schwarzschild (**416**, 25) in another of the fundamental papers of radiative transfer theory. Schwarzschild was able to demonstrate convincingly that the dominant mode of energy transport in the photosphere of the sun is radiative. Since his work, a number of analyses of radiative stability have been carried out for a variety of stellar types; results are summarized in (**638**, 215; **11**, 449; **654**, 432). The basic picture that emerges for a star like the sun is that radiative equilibrium obtains to continuum optical depths of order unity, and that below this depth the atmosphere becomes unstable against convection. Convection zones exist below the outer radiative zone in all stars of spectral type later than about F5. For earlier spectral types, radiative equilibrium prevails throughout the entire outer envelope of a star. In this book primary concern will be given to the early-type stars, and accordingly

emphasis will be given to the radiative equilibrium regime. The theory of convective transport is not, at present, in a fully satisfactory form, and only the *mixing-length theory* (a phenomenological approach that has been widely applied in astrophysics) will be described (§7-3).

Let us now consider some of the implications of radiative equilibrium; assume the medium is static and the radiation field is time-independent. From the discussion in §2-1, it is clear that the total energy removed from the beam is

$$\int_0^\infty d\nu \oint d\omega\, \chi(\mathbf{r}, \nu) I(\mathbf{r}, \mathbf{n}, \nu) = 4\pi \int_0^\infty \chi(\mathbf{r}, \nu) J(\mathbf{r}, \nu)\, d\nu \qquad (2\text{-}81)$$

where $\chi$ denotes the total extinction coefficient. The total energy delivered by the material to the radiation field is

$$\int_0^\infty d\nu \oint d\omega\, \eta(\mathbf{r}, \nu) = 4\pi \int_0^\infty \chi(\mathbf{r}, \nu) S(\mathbf{r}, \nu)\, d\nu \qquad (2\text{-}82)$$

where equation (2-35) has been used. The condition of radiative equilibrium demands that the total energy absorbed by a given volume of material must equal the total energy emitted; thus at each point in the atmosphere

$$4\pi \int_0^\infty [\eta(\mathbf{r}, \nu) - \chi(\mathbf{r}, \nu) J(\mathbf{r}, \nu)]\, d\nu = 0 \qquad (2\text{-}83a)$$

or

$$4\pi \int_0^\infty \chi(\mathbf{r}, \nu)[S(\mathbf{r}, \nu) - J(\mathbf{r}, \nu)]\, d\nu = 0 \qquad (2\text{-}83b)$$

*Exercise 2-13*: Suppose that $S_\nu$ is given by equation (2-39). Show that, in the condition of radiative equilibrium, the scattering terms cancel out to yield

$$\int_0^\infty \kappa_\nu B_\nu(T)\, d\nu = \int_0^\infty \kappa_\nu J_\nu\, d\nu$$

Using equation (2-83) in equation (2-69) we have, alternatively,

$$\nabla \cdot \mathscr{F} \equiv 0 \qquad (2\text{-}84)$$

Hence *in planar geometry the condition of radiative equilibrium is equivalent to the requirement that the depth derivative of the flux is zero*—i.e., *the flux is constant*. Physically, equations (2-83) and (2-84) have the same meaning, but mathematically the requirement $\mathscr{F} =$ constant is rather different from the expressions in equations (2-83); either form of the constraint may be used in constructing model atmospheres.

Because the total flux is constant in a planar atmosphere, it may be used as a parameter that describes the atmosphere; an equivalent quantity often employed is the *effective temperature*. From Exercise 1-6 we know that the integrated flux from a black body of absolute temperature $T$ is $\mathscr{F}_{BB} = \sigma_R T^4$.



Although the radiation emerging from a star is by no means Planckian, it is nevertheless customary to define the effective temperature as the temperature a black body would have in order to emit the actual stellar flux—i.e.,

$$\sigma_R T_{\text{eff}}^4 \equiv \int_0^\infty \mathscr{F}_\nu\, d\nu = 4\pi \int_0^\infty H_\nu\, d\nu = L/(4\pi R^2) \qquad (2\text{-}85)$$

Here $L$ is the total luminosity and $R$ is the radius of the star; the atmospheric thickness is assumed to be negligible compared to $R$. Although $T_{\text{eff}}$ has only an indirect physical significance, it is a convenient parameter with which to characterize the atmosphere, for typically the actual kinetic temperature $T$ will equal $T_{\text{eff}}$ near the depth from which the continuum radiation emerges (i.e., unit optical depth at frequencies where the opacity is lowest). In *spherical geometry* equation (2-84) implies

$$r^2 \mathscr{F} = \text{constant} = L/4\pi \qquad (2\text{-}86)$$

In an extended atmosphere it is no longer really possible to choose a unique radius $R$ for the star and to define a unique value of $T_{\text{eff}}$; rather, $L$ or the quantity $r^2\mathscr{F}$ should be regarded as fundamental. However the identification $R = r(\tau_R = \frac{2}{3})$ is sometimes made, and a value of "$T_{\text{eff}}$" derived with this radius; here $\tau_R$ is the *Rosseland optical depth* scale (cf. §3-2).

Finally, it is important to return to the point raised in the discussion of the formal solution. Suppose that the opacity $\kappa_\nu$ is independent of $T$; then for physically reasonable distributions of $\kappa_\nu$, the integral $\int \kappa_\nu B_\nu(\tau)\, d\nu$ (which gives the total thermal emission) will be a monotone increasing function of $T$. Thus when we fix the total thermal emission at some value, we fix the local value of $T$. It is then clear from equation (2-83b) (and from the result of Exercise 2-13) that the *local* value of $T$ is determined by the mean intensity, which depends upon the *global* properties of the atmosphere because it follows from a solution of the transfer equation. Thus the temperature at a *given* point in the atmosphere is to some extent determined by the temperature at *all other* points and, at the same time, helps to establish the temperature structure elsewhere. This *nonlocalness* in the problem is a result of radiative transfer, through which photons moving from one point in the medium to another lead to a fundamental coupling (i.e., interdependence) of the properties at those points.

## 2-5 The Diffusion Approximation

At great depths in a semi-infinite atmosphere the properties of the radiation field and the nature of the transfer equation become extremely simple. We can obtain in a straightforward way an asymptotic solution that applies

throughout the interior of a star (except, of course, in convective zones), and that provides a lower boundary condition on the transfer problem in the stellar atmosphere. Consider first the properties of the radiation field. At depths in the medium much larger than a photon mean-free-path, the radiation is effectively trapped, becomes essentially isotropic, and (eventually) approaches thermal equilibrium so that $S_\nu \to B_\nu$. Choose a reference point $\tau_\nu \gg 1$, and expand $S_\nu$ as a power-series:

$$S_\nu(t_\nu) = \sum_{n=0}^{\infty} [d^n B_\nu/d\tau_\nu{}^n](t_\nu - \tau_\nu)^n/n! \tag{2-87}$$

Calculating the specific intensity for $0 \leqslant \mu \leqslant 1$ from this source function with equation (2-50) we have

$$I_\nu(\tau_\nu, \mu) = \sum_{n=0}^{\infty} \mu^n \frac{d^n B_\nu}{d\tau_\nu{}^n} = B_\nu(\tau_\nu) + \mu \frac{dB_\nu}{d\tau_\nu} + \mu^2 \frac{d^2 B_\nu}{d\tau_\nu{}^2} + \cdots \tag{2-88}$$

A similar result for $-1 \leqslant \mu \leqslant 0$ follows from equation (2-51) and differs from equation (2-88) only by terms of order $e^{-\tau/\mu}$; in the limit of great depth the latter vanish and equation (2-88) applies on the full range $-1 \leqslant \mu \leqslant 1$. By substitution of equation (2-88) into the appropriate definitions we find for the moments

$$J_\nu(\tau_\nu) = \sum_{n=0}^{\infty} (2n+1)^{-1}(d^{2n}B_\nu/d\tau_\nu{}^{2n})$$

$$= B_\nu(\tau_\nu) + \frac{1}{3}(d^2 B_\nu/d\tau_\nu{}^2) + \cdots \tag{2-89a}$$

$$H_\nu(\tau_\nu) = \sum_{n=0}^{\infty} (2n+3)^{-1}(d^{2n+1}B_\nu/d\tau_\nu^{2n+1})$$

$$= \frac{1}{3}(dB_\nu/d\tau_\nu) + \cdots \tag{2-89b}$$

and

$$K_\nu(\tau_\nu) = \sum_{n=0}^{\infty} (2n+3)^{-1}(d^{2n}B_\nu/d\tau_\nu{}^{2n})$$

$$= \frac{1}{3}B_\nu(\tau_\nu) + \frac{1}{5}(d^2 B_\nu/d\tau_\nu{}^2) + \cdots \tag{2-89c}$$

Note that only even-order terms survive in the even moments $J_\nu$ and $K_\nu$ and only odd-order terms in $H_\nu$.

We now inquire how rapidly these series converge. The *derivatives* can be approximated, at least to order-of-magnitude, by appropriate *differences*—



i.e., $|d^n B_\nu/d\tau_\nu{}^n| \sim B_\nu/\tau_\nu{}^n$. Then it is clear that the ratio of successive terms in the series is of order $O(1/\tau_\nu{}^2)$ or $O(1/\langle\chi_\nu\rangle^2\,\Delta z^2)$ where $\langle\chi_\nu\rangle$ is the average opacity over the path-length $\Delta z$. In terms of the photon mean-free-path $l_\nu \sim 1/\chi_\nu$, the convergence factor is $O(l_\nu{}^2/\Delta z^2)$. It is clear that the convergence is quite rapid; indeed the estimate just given turns out to be conservative. Also, it is obvious that convergence will be most rapid at frequencies where the material is quite opaque. For a star as a whole one expects $\Delta z$ to be some significant fraction of a stellar radius, say $\Delta z \sim 10^{10}$ cm, while $\langle\chi\rangle \sim 1$ (which implies a photon mean-free-path of 1 cm), so that the convergence factor of the series is of the order $10^{-20}$. It is clear that in the deep interior of the star only the leading terms are required.

In the limit of large depth we may therefore write

$$I_\nu(\tau_\nu, \mu) \approx B_\nu(\tau_\nu) + \mu(dB_\nu/d\tau_\nu) \tag{2-90a}$$

$$J_\nu(\tau_\nu) \approx B_\nu(\tau_\nu) \tag{2-90b}$$

$$H_\nu(\tau_\nu) \approx \frac{1}{3}(dB_\nu/d\tau_\nu) \tag{2-90c}$$

$$K_\nu(\tau_\nu) \approx \frac{1}{3}B_\nu(\tau_\nu) \tag{2-90d}$$

In equation (2-90a) we have retained two terms so as to account for the nonzero flux [cf. equation (2-90c)]. Note that equations (2-90b) and (2-90d) show that $\lim_{\tau\to\infty}[K_\nu(\tau_\nu)/J_\nu(\tau_\nu)] = \frac{1}{3}$, which is what we would expect for isotropic radiation; we shall show below that the ratio of the anisotropic to isotropic term in $I(\tau, \mu)$ becomes vanishingly small as $\tau \to \infty$, so that the limit just found is appropriate. Insofar as the specific intensity as given by equation (2-90a) was computed from the formal solution of the transfer equation, in effect using a source function $S_\nu(t_\nu) = B_\nu(\tau_\nu) + (t_\nu - \tau_\nu)(dB_\nu/d\tau_\nu)$, equations (2-90) should obviate further use of the transfer equation. It is easy to verify that this is so, for by inspection one sees that substitution from equations (2-90) reduces both the transfer equation (2-36) and the zero-order moment equation (2-71) to the single requirement $d^2 B_\nu/d\tau_\nu{}^2 = 0$ (already assumed), while the first-order moment equation (2-79) is identical to (2-90c). Thus, in effect, at great depth the transfer problem reduces to the single equation

$$H_\nu = \frac{1}{3}\left(\frac{\partial B_\nu}{\partial \tau_\nu}\right) = -\frac{1}{3}\left(\frac{1}{\chi_\nu}\frac{\partial B_\nu}{\partial T}\right)\left(\frac{dT}{dz}\right) \tag{2-91}$$

It is clear that equations (2-90) and (2-91) can be used, as mentioned in §2-2, to set lower boundary conditions on the transfer equation in a semi-infinite atmosphere.

Equation (2-91) [and also equations (2-90)] is referred to as the *diffusion approximation*, primarily because of its formal similarity to other diffusion equations, which are of the form

(flux) = (diffusion coefficient) × (gradient of relevant physical variable)

e.g., $\mathbf{\Phi} = -\kappa\,\nabla T$ for heat conduction. The coefficient $\frac{1}{3}\chi_\nu^{-1}(\partial B_\nu/\partial T)$ is, in fact, sometimes called the *radiative conductivity*, a designation that is quite appropriate in view of the fact that $\chi_\nu^{-1} = l_\nu$ is the photon mean-free-path. Note that equation (2-91) exhibits the essential physical content of our earlier result that the flux computed by application of the $\Phi$-operator to a linear source function depends only on the gradient of $S$ [see discussion following equation (2-65)]. Also it shows that *the mere fact that energy emerges from the star implies that the temperature must increase inward.* Indeed, replacing $H$ with $(L/4\pi R^2)$ and $(dB/dz)$ with $(\sigma_R T_c^4/\pi R)$, and taking $\langle\chi\rangle \approx 1$, it is easy to show that the central temperature $T_c$ of the sun must be of the order of $6 \times 10^6$ °K, a result consistent with our earlier statement that the ultimate energy source in a star is thermonuclear energy-release at the center.

In an intuitive picture of diffusion, one usually conceives of a slow leakage from a reservoir of large capacity by means of a seeping action. These ideas apply in the radiative diffusion limit as well. The diffusion approximation becomes valid at great optical depth (i.e., many photon mean-free-paths from the surface) whence many individual photon flights, with successive absorptions and emissions, are required before the photon finally trickles to the surface, and issues forth into interstellar space.

If we integrate equations (2-90a–c) over all frequencies, we obtain $I(\tau, \mu) \approx B(\tau) + 3\mu H$. The ratio of the anisotropic to isotropic terms gives a measure of the "drift" in the radiation flow; this ratio is

$$\frac{\text{Anisotropic term}}{\text{Isotropic term}} \sim \frac{3H}{B} = \frac{3}{4}\left(\frac{\sigma_R T_{\text{eff}}^4}{\pi}\right)\Big/\left(\frac{\sigma_R T^4}{\pi}\right) \sim \left(\frac{T_{\text{eff}}}{T}\right)^4 \qquad (2\text{-}92)$$

Clearly at great depth, where $T$ becomes $\gg T_{\text{eff}}$, the "leak" becomes ever smaller. The same result is found by a physical argument from a slightly different point of view. If $\pi F$ is the energy flux carried from an element of material by photons of velocity $c$, the rate of *energy flow* per unit volume is $(\pi F/c)$; the *energy content* per unit volume is $(4\pi J/c) \approx (4\pi B/c)$, so that (Rate of energy flow)/(Energy content) $= (F/4B) = \frac{1}{4}(T_{\text{eff}}/T)^4$. Again, we see that diffusion, in the intuitive sense described above, occurs at great depth where $T \gg T_{\text{eff}}$, while *free flow* of radiant energy occurs at the surface where $T \approx T_{\text{eff}}$.

# 3

# *The Grey Atmosphere*

The grey atmosphere problem provides an excellent introduction to the study of radiative transfer in stellar atmospheres. The nature of the defining assumptions is such that the problem becomes independent of the physical state of the material, and requires the solution of a relatively simple transfer equation. At the same time, the grey problem demonstrates how the constraint of radiative equilibrium can be satisfied, and the solution can be related to more general and more realistic physical situations. Furthermore, an exact solution of the problem can be obtained, and this provides a comparison standard against which we can evaluate the worth of various approximate numerical methods that can be applied in more complex cases.

## 3-1 Statement of the Problem

The problem is posed by making the simplifying assumption that the opacity of the material is independent of frequency; i.e., $\chi_\nu \equiv \chi$. This assumption is of course unrealistic in many cases. Yet as we shall see in later chapters, the opacity in some stars (e.g., the sun) is not too far from being grey and, in

addition, it is possible partially to reduce the nongrey problem to the grey case by suitable choices of *mean opacities*. Thus the solution also provides a valuable starting approximation in the analysis of nongrey atmospheres.

If we assume $\chi_\nu \equiv \chi$, then the standard planar transfer equation (2-36) becomes

$$\mu(\partial I_\nu/\partial\tau) = I_\nu - S_\nu \tag{3-1}$$

Then by integrating over frequency, and writing

$$I \equiv \int_0^\infty I_\nu \, d\nu \tag{3-2}$$

and similarly for $J$, $S$, $B$, etc., we have

$$\mu(\partial I/\partial\tau) = I - S \tag{3-3}$$

If we impose the constraint of radiative equilibrium [equation 2-83b)], we require

$$\int_0^\infty \chi J_\nu \, d\nu = \int_0^\infty \chi S_\nu \, d\nu \tag{3-4}$$

which, for grey material, reduces to $J \equiv S$. Thus equation (3-3) becomes

$$\mu(\partial I/\partial\tau) = I - J \tag{3-5}$$

which has the formal solution [equation (2-57)]

$$J(\tau) = \Lambda_\tau[S(t)] = \Lambda_\tau[J(t)] = \frac{1}{2}\int_0^\infty J(t)E_1|t - \tau| \, dt \tag{3-6}$$

Equation (3-6) is a linear integral equation for $J$ known as *Milne's equation*; the grey problem itself is sometimes called *Milne's problem*. It is important to recognize that, when a solution of equation (3-6) is obtained, it satisfies *simultaneously* the transfer equation and the constraint of radiative equilibrium. The determination of such solutions in the nongrey case will occupy most of Chapter 7.

If we now introduce the additional hypothesis of LTE, then $S_\nu \equiv B_\nu(T)$—which, from the condition of radiative equilibrium, implies that

$$J(\tau) = S(\tau) = B[T(\tau)] = (\sigma_R T^4)/\pi \tag{3-7}$$

Thus, if we are given $J(\tau)$, the solution of the integral equation (3-6), then the additional premise of LTE allows us to associate a temperature with the radiative equilibrium radiation field via equation (3-7).

Several important results may be obtained from moments of equation (3-5). Taking the zero-order moment and imposing radiative equilibrium we have

$$(dH/d\tau) = J - S = J - J \equiv 0 \tag{3-8}$$



which implies *the flux is constant*, while the first moment gives

$$(dK/d\tau) = H \tag{3-9}$$

which, because $H$ is constant, yields the *exact integral*

$$K(\tau) = H\tau + c = \frac{1}{4}F\tau + c \tag{3-10}$$

To make further progress, we must relate $J(\tau)$ to $K(\tau)$. This is easily done on the basis of the discussion in §2-5, where we showed that at great depth the specific intensity is quite accurately represented by $I(\mu) = I_0 + I_1\mu$, which produces a nonzero flux and also implies that, for $\tau \gg 1$, $K(\tau) = \frac{1}{3}J(\tau)$. Thus the fact that $K(\tau) \to \frac{1}{4}F\tau$ for $\tau \gg 1$ implies that at great depth

$$J(\tau) \to \frac{3}{4}F\tau \quad (\tau \gg 1) \tag{3-11}$$

That is, asymptotically the mean intensity varies linearly with optical depth. On general grounds we expect the behavior of $J(\tau)$ to depart most from linearity at the surface [note equation (2-63)], which suggests that a reasonable general expression for $J(\tau)$ is

$$J(\tau) = \frac{3}{4}F[\tau + q(\tau)] = \frac{3}{4}(\sigma_R T_{\text{eff}}^4/\pi)[\tau + q(\tau)] \tag{3-12}$$

The function $q(\tau)$, known as the *Hopf function*, remains to be determined; from equation (3-6) it is clear that $q(\tau)$ is a solution of the equation

$$\tau + q(\tau) = \frac{1}{2}\int_0^\infty [t + q(t)]E_1|t - \tau| \, dt \tag{3-13}$$

Finally, we notice that because

$$\lim_{\tau\to\infty}\left[\frac{1}{3}J(\tau) - K(\tau)\right] = \frac{1}{4}F\lim_{\tau\to\infty}[\tau + q(\tau) - \tau - c] \equiv 0 \tag{3-14}$$

we have $c = q(\infty)$ and hence can write equation (3-10) as

$$K(\tau) = \frac{1}{4}F[\tau + q(\infty)] \tag{3-15}$$

The solution of the grey problem consists of the specification of $q(\tau)$. Given $q(\tau)$, the temperature distribution is obtained by combining equations (3-7) and (3-12) into the relation

$$T^4 = \frac{3}{4}T_{\text{eff}}^4[\tau + q(\tau)] \tag{3-16}$$

We shall derive approximate expressions for $q(\tau)$ in §3-3 and describe the exact solution in §3-4. First, however, it is useful to delineate the nature, and extent, of the correspondence between the grey and nongrey problems.

## 3-2 Relation to the Nongrey Problem: Mean Opacities

The opacity in real stellar atmospheres usually exhibits strong frequency variations, at least when spectral lines are present. Although numerical methods now exist that allow a refined solution of nongrey transfer equations and an accurate determination of the temperature structure in a nongrey atmosphere, the calculations are, at best, laborious, and it is important to ask whether a significant connection exists between the grey and nongrey cases. We shall show in this section that such a connection, though limited in scope, does exist, and that, among other things, it permits the temperature distribution of the deepest atmospheric layers to be determined quite accurately from the grey solution.

Let us first compare side-by-side the grey and nongrey transfer equations. Starting with the transfer equation and calculating the zero and first-order moments we have, in the nongrey and grey cases respectively:

$$\mu(\partial I_\nu/\partial z) = \chi_\nu(S_\nu - I_\nu) \quad (3\text{-}17a) \qquad \mu(\partial I/\partial z) = \chi(J - I) \quad (3\text{-}17b)$$

$$(\partial H_\nu/\partial z) = \chi_\nu(S_\nu - J_\nu) \quad (3\text{-}18a) \qquad (dH/dz) = 0 \quad (3\text{-}18b)$$

$$(\partial K_\nu/\partial z) = -\chi_\nu H_\nu \quad (3\text{-}19a) \qquad (dK/dz) = -\chi H \quad (3\text{-}19b)$$

Here variables without frequency subscripts denote *integrated* quantities, as in equation (3-2). We now ask whether it is possible to define a mean opacity $\bar{\chi}$, formed as a weighted average of the monochromatic opacity, in such a way that the monochromatic transfer equation, or one of its moments, when integrated over frequency, has exactly the same form as the grey equation. Several possible definitions have been suggested.

### FLUX–WEIGHTED MEAN

Suppose we wish to define a mean opacity in such a way as to guarantee an exact correspondence between the integrated form of equation (3-19a) and the grey equation (3-19b). If such a mean can be constructed, then the relation $K(\bar{\tau}) = H\bar{\tau} + c$ will again be an exact integral, as it was in the grey case. Integrating equation (3-19a) over all frequencies we have

$$-(dK/dz) = \int_0^\infty \chi_\nu H_\nu \, d\nu = \bar{\chi}_F H \tag{3-20}$$



where the second equality produces the desired identification with equation (3-19b) if we define

$$\bar{\chi}_F \equiv H^{-1} \int_0^\infty \chi_\nu H_\nu \, d\nu \tag{3-21}$$

The opacity $\bar{\chi}_F$ is called the *flux-weighted* mean. Note that this choice does *not* reduce the nongrey problem completely to the grey problem, for the monochromatic equation (3-18a) does not transform into equation (3-18b) with this choice of $\bar{\chi}$. Further, there is the practical problem that $H_\nu$ is not known a priori, and therefore $\bar{\chi}_F$ cannot actually be calculated until after the transfer equation is solved. This latter difficulty can be overcome by an iteration between construction of models and calculation of $\bar{\chi}_F$. Although the desired goal has not been fully attained, the fact that the flux-weighted mean preserves the $K$-integral is important, for it implies that the correct value is recovered for the *radiation pressure* [cf. equation (1-41)]. It also follows that the correct value of the *radiation force*, which is the gradient of the radiation pressure, is likewise obtained. Thus from equation (2-77) we have

$$(dp_R/d\bar{\tau}) = -\bar{\chi}_F{}^{-1}(dp_R/dz) = (4\pi/c\bar{\chi}_F) \int_0^\infty \chi_\nu H_\nu \, d\nu = (4\pi H/c) = (\sigma T_{\text{eff}}^4/c) \tag{3-22}$$

so that use of the flux-mean opacity produces a simple expression for the radiation pressure gradient. This is a result of practical value in the computation of model atmospheres for early-type stars, because in these objects radiation forces strongly affect the pressure (or density) structure of the atmosphere via the condition of hydrostatic equilibrium (or momentum balance in steady flow).

### ROSSELAND MEAN

Alternatively, suppose we wish to obtain the correct value for the integrated energy flux. From equations (3-19) it follows that this may be done if $\bar{\chi}$ is chosen such that

$$-\int_0^\infty \chi_\nu{}^{-1}(\partial K_\nu/\partial z) \, d\nu = \int_0^\infty H_\nu \, d\nu = H \equiv -\bar{\chi}^{-1}(dK/dz) \tag{3-23}$$

or, equivalently,

$$\bar{\chi}^{-1} \equiv \int_0^\infty \chi_\nu{}^{-1}(\partial K_\nu/\partial z) \, d\nu \Big/ \int_0^\infty (\partial K_\nu/\partial z) \, d\nu \tag{3-24}$$

Again we face the practical difficulty that $K_\nu$ is not known a priori, and hence the indicated calculation cannot be performed until the transfer equation is

solved. But the mean defined in equation (3-24) can be approximated in the following way: at great depth in the atmosphere, $K_\nu \to \frac{1}{3}J_\nu$ while $J_\nu \to B_\nu$. Thus may write $(\partial K_\nu/\partial z) \approx \frac{1}{3}(\partial B_\nu/\partial T)(dT/dz)$. We then define the mean opacity $\bar{\chi}_R$ as

$$\frac{1}{\bar{\chi}_R} \equiv \frac{\frac{1}{3}\left(\frac{dT}{dz}\right)\int_0^\infty \left(\frac{1}{\chi_\nu}\right)\frac{\partial B_\nu}{\partial T}\,d\nu}{\frac{1}{3}\left(\frac{dT}{dz}\right)\int_0^\infty \frac{\partial B_\nu}{\partial T}\,d\nu} = \frac{\int_0^\infty \frac{1}{\chi_\nu}\frac{\partial B_\nu}{\partial T}\,d\nu}{\int_0^\infty \frac{\partial B_\nu}{\partial T}\,d\nu} \tag{3-25}$$

or

$$\bar{\chi}_R{}^{-1} \equiv (\pi/4\sigma_R T^3)\int_0^\infty \chi_\nu{}^{-1}(\partial B_\nu/\partial T)\,d\nu \tag{3-26}$$

The opacity $\bar{\chi}_R$ is called the *Rosseland mean* in honor of its originator. Note that the harmonic nature of the averaging process gives highest weight to those regions where the opacity is lowest, and, where as a result, the greatest amount of radiation is transported—a very desirable feature. Again the use of $\bar{\chi}_R$ or the mean defined by equation (3-24) does not permit a correspondence between equations (3-18a) and (3-18b), and hence does not allow the nongrey problem to be replaced by the grey problem. On the other hand, it is obvious that the approximations made to obtain equation (3-26) are precisely those introduced in the derivation of the diffusion approximation to the transfer equation (2-91); i.e.,

$$H_\nu = -\frac{1}{3}\left(\frac{1}{\chi_\nu}\frac{\partial B_\nu}{\partial T}\right)\left(\frac{dT}{dz}\right)$$

Hence use of $\bar{\chi}_R$ is consistent with the diffusion approximation. Therefore on the Rosseland-mean optical-depth scale $\bar{\tau}_R$ we must recover the correct asymptotic solution of the transfer equation and the correct flux transport at great depth. This implies that at great depth ($\bar{\tau}_R \gg 1$) the temperature structure is quite accurately given by the relation $T^4 = \frac{3}{4}T_{\text{eff}}^4[\bar{\tau}_R + q(\bar{\tau}_R)]$; see equation (3-16). It is therefore clear why Rosseland mean opacities are employed in studies of stellar interiors. Note also that so long as the diffusion approximation is valid, a simple expression can be written for the radiation force, namely

$$(dp_R/d\bar{\tau}_R) = (16\pi\sigma_R T^3/3c\bar{\chi}_R)(-dT/dz) \tag{3-27}$$

*Exercise 3-1*: Derive equation (3-27).

Although the diffussion approximation is nearly exact at great depth, and provides the very useful results just discussed, it *must* of course break down at the surface, and exact flux conservation is *not* guaranteed in the upper layers by use of the Rosseland mean, nor will it give the temperature structure



or the radiation force correctly in the outermost regions of the atmosphere. This point must be recognized clearly, for it is precisely these layers in which spectrum-formation occurs, and hence which are of primary interest in the analysis of stellar spectra.

### PLANCK AND ABSORPTION MEANS

Several other expressions for mean opacities may be chosen. For example, if we demand that the mean be defined to yield the correct integrated value of the thermal emission, then we require

$$\bar{\kappa}_P \equiv \left[\int_0^\infty \kappa_\nu B_\nu(T)\,d\nu\right] \Big/ B(T) = \pi\int_0^\infty \kappa_\nu B_\nu(T)\,d\nu \Big/ \sigma_R T^4 \tag{3-28}$$

Note that only the true absorption is used, and scattering is omitted. The opacity $\bar{\kappa}_P$ is known as the *Planck mean*; it has the advantage of being calculable without having to solve the transfer equation. On the other hand, $\bar{\kappa}_P$ does not permit a reduction of equation (3-18a) to (3-18b) nor of (3-19a) to (3-19b), and therefore it lacks the desirable features possessed by $\bar{\chi}_F$ or $\bar{\chi}_R$. Nevertheless this mean does have additional significance.

In particular, near the surface of the star, the physical content of the condition of radiative equilibrium is expressed most directly by equation (3-4). In view of this constraint, a correspondence between equations (3-18a) and (3-18b) can be made near the surface if $\bar{\kappa}$ satisfies the relation

$$\int_0^\infty (\kappa_\nu - \bar{\kappa})(J_\nu - B_\nu)\,d\nu = 0 \tag{3-29}$$

Once the material becomes optically thin (i.e., $\tau_\nu < 1$ at all frequencies), $J_\nu$ becomes essentially fixed, and the integral above will be dominated by those frequencies where $\kappa_\nu \gg \bar{\kappa}$. If we represent $B_\nu$ by a linear expansion on a $\bar{\tau}$-scale, i.e.,

$$B_\nu(t) = B_\nu(\bar{\tau}) + (dB_\nu/d\bar{\tau})(t - \bar{\tau}) \approx B_\nu(\bar{\tau}) + (dB_\nu/d\bar{\tau})(\bar{\kappa}/\kappa_\nu)(t_\nu - \tau_\nu),$$

then by application of the $\Lambda$-operator we find [cf. equations (2-57), (2-58), and (2-63)],

$$J_\nu(\bar{\tau}) - B_\nu(\bar{\tau}) \approx -\frac{1}{2}B_\nu(\bar{\tau})E_2(\tau_\nu) + (\bar{\kappa}/\kappa_\nu)(dB_\nu/d\bar{\tau})\left[\frac{1}{2}E_3(\tau_\nu) + \frac{1}{2}\tau_\nu E_2(\tau_\nu)\right] \tag{3-30}$$

In the limit $\tau \to 0$, $E_2(\tau) \to 1$, and $E_3(\tau) \to \frac{1}{2}$, so the first term yields $-\frac{1}{2}B_\nu(\bar{\tau})$ while the second becomes *least* important when $\kappa_\nu \gg \bar{\kappa}$ [precisely the region of highest weight in equation (3-29)]. Thus to satisfy equation (3-29), $\bar{\kappa}$

should essentially fulfill the requirement

$$\int_0^\infty \kappa_\nu B_\nu \, d\nu = \bar{\kappa} \int_0^\infty B_\nu \, d\nu \tag{3-31}$$

which shows that the Planck mean is the choice most nearly consistent with the requirement of radiative equilibrium near the surface.

Alternatively, we might demand that the mean opacity yield the correct total for the amount of energy absorbed. We then obtain the *absorption mean*

$$\bar{\kappa}_J \equiv \int_0^\infty \kappa_\nu J_\nu \, d\nu \Big/ J \tag{3-32}$$

Again only the true absorption is included, and scattering is omitted. As was true for $\bar{\chi}_F$, we cannot calculate $\bar{\kappa}_J$ until a solution of the transfer equation has been obtained. Further, $\bar{\kappa}_J$ does not permit a strict correspondence between the grey and nongrey forms of the transfer equation or any of its moments (as was also true for the Planck mean).

SUMMARY

We have seen that no one of the mean opacities described above allows, in itself, a complete correspondence of the nongrey problem to the grey problem. Yet mean opacities provide a useful first estimate of the temperature structure in a stellar atmosphere if we assume, as a starting approximation, $T(\bar{\tau}_R) = T_{\text{grey}}(\bar{\tau}_R)$, and then improve this estimate with a *correction procedure* that is designed to enforce radiative equilibrium for the nongrey radiation field. Indeed, the mean opacities $\bar{\chi}_F$, $\bar{\kappa}_P$, and $\bar{\kappa}_J$ appear explicitly in some temperature-correction procedures.

From an historical point of view, it should be recognized that, before the advent of high-speed computers, the nongrey atmosphere problem required far too much calculation to permit a direct attack, and the use of $\bar{\chi}_R$ and $\bar{\kappa}_P$ provided a practical method of approaching an otherwise intractable problem. In fact, the answers obtained in this way often do not compare too unfavorably with more recent results despite the apparent crudeness of the approximation. Only some of the more basic properties of mean opacities have been mentioned here; further information may be found in (**419**) and (**361**, §§34–35).

## 3-3 Approximate Solutions

THE EDDINGTON APPROXIMATION

In §2-5 it was shown that, at great depth in a stellar atmosphere, the relation $J = 3K$ holds; further, in §1-4 (cf. Exercise 1-13), it was demonstrated



that this relation is also valid for a variety of other situations, including the two-stream approximation, which provides a rough representation of the radiation field near the boundary. In view of these results, Eddington made the simplifying assumption that $J \equiv 3K$ *everywhere* in the atmosphere. Then the exact integral $K = \frac{1}{4}F\tau + c$ implies that in the Eddington approximation $J_E(\tau) = \frac{3}{4}F\tau + c'$. To evaluate the constant $c'$ we calculate the emergent flux and fit it to the desired value. Thus from equation (2-59) we have

$$F(0) = 2\int_0^\infty \left(\frac{3}{4}F\tau + c'\right) E_2(\tau)\, d\tau = 2c'E_3(0) + \frac{3}{4}F\left[\frac{4}{3} - 2E_4(0)\right] \tag{3-33}$$

so that, using the relation $E_n(0) = 1/(n-1)$ and demanding $F(0) \equiv F$, we find $c' = \frac{1}{2}F$. Thus

$$J_E(\tau) = \frac{3}{4}F\left(\tau + \frac{2}{3}\right) \tag{3-34}$$

In Eddington's approximation $q(\tau) \equiv \frac{2}{3}$. Imposing the constraint of radiative equilibrium and the assumption of LTE we have from equation (3-16)

$$T^4 \approx \frac{3}{4}T_{\text{eff}}^4\left(\tau + \frac{2}{3}\right) \tag{3-35}$$

Equation (3-35) predicts that the ratio of the boundary temperature to the effective temperature is $(T_0/T_{\text{eff}}) = (\frac{1}{2})^{1/4} = 0.841$, which agrees fairly well with the *exact* value $(T_0/T_{\text{eff}}) = (3^{1/2}/4)^{1/4} = 0.8114$. Assuming $S(\tau) = J_E(\tau)$ we may calculate the angular dependence of the emergent radiation field in the Eddington approximation by substituting equation (3-34) into equation (2-52) to obtain

$$I_E(0, \mu) = \frac{3}{4}F\int_0^\infty \left(\tau + \frac{2}{3}\right)\mu^{-1}\exp(-\tau/\mu)\, d\tau = \frac{3}{4}F\left(\mu + \frac{2}{3}\right) \tag{3-36}$$

which yields a very specific form for the Eddington–Barbier relation [cf. equation (2-53)]. The *center* of a star's disk, as seen by an external observer, corresponds to $\theta = 0°$, or $\mu = 1$; the *limb* is at $\theta = 90°$, $\mu = 0$. The ratio $I(0, \mu)/I(0, 1)$, which gives the intensity at angle $\theta = \cos^{-1}\mu$ relative to disk-center, is referred to as the *limb-darkening law*. In the Eddington approximation the limb-darkening is

$$I_E(0, \mu)/I_E(0, 1) = \frac{3}{5}\left(\mu + \frac{2}{3}\right) \tag{3-37}$$

This result predicts the limb intensity to be 40 percent of the central intensity. Observations of the sun in the visual regions of the spectrum are actually in good agreement with this value and, in fact, it was precisely this agreement

that led K. Schwarzschild (**416**, 25) to propose the validity of radiative equilibrium in the outermost layers of the solar photosphere.

Equation (3-35) predicts that $T = T_{\text{eff}}$ when $\tau = \frac{2}{3}$. This result has given rise to the useful conceptual notion that the "effective depth" of continuum formation is $\tau \approx \frac{2}{3}$; in fact, this is often a rather good estimate. In support of this idea we may note that a photon emitted outward from $\tau = \frac{2}{3}$ has a chance of the order of $e^{-0.67} \approx 0.5$ of emerging from the surface; this corresponds in a reasonable way with the place we would intuitively identify with the region of continuum formation.

*Exercise 3-2*: The Eddington–Barbier relation shows that the intensity $I(0, \mu)$ is characteristic of $S(\tau)$ at $\tau(\mu) \approx \mu$. Show then that the average depth that characterizes the *flux* is $\langle\tau\rangle = \frac{2}{3}$.

Anticipating the exact solution given in Table 3-2, we can evaluate the accuracy of $J_E(\tau)$; one finds that the worst error occurs at the surface, where $\Delta J/J = (J_E - J_{\text{exact}})/J_{\text{exact}} = 0.155$. Both the size of the error and the fact that it occurs at $\tau = 0$ are unsurprising when we recognize that the basic assumption upon which the derivation is based, namely $J \equiv 3K$, is known to be inaccurate at the surface. We know that $J(\tau)$ must satisfy the integral equation (3-6), and we know further that the $\Lambda$-operator has its largest effect at $\tau = 0$ [see equation (2-63) and associated discussion]. This suggests that an improved estimate of $J(\tau)$ can be obtained from

$$J_E^{(2)}(\tau) = \Lambda_\tau[J_E(t)] = \Lambda_\tau\left[\frac{3}{4}F\left(t + \frac{2}{3}\right)\right] = \frac{3}{4}F\left[\tau + \frac{2}{3} - \frac{1}{3}E_2(\tau) + \frac{1}{2}E_3(\tau)\right] \tag{3-38}$$

Recalling the properties of $E_n(\tau)$, it is clear that the largest difference between $J_E^{(2)}(\tau)$ and $J_E(\tau)$ occurs at the surface, where we find $J_E^{(2)}(0)/J_E(0) = \frac{7}{8}$. The new estimate of $T_0/T_{\text{eff}}$ is thus $(\frac{7}{16})^{1/4} = 0.813$ (exact value is 0.8114) and $q(0)$ drops from $\frac{2}{3}$ to $\frac{7}{12} = 0.583$ (exact value is $1/\sqrt{3} = 0.577$).

It is thus clear that an application of the $\Lambda$-operator has produced a dramatic improvement in the solution near the surface. Note, however, that there is *no* improvement in the solution at $\tau \to \infty$, where $q$ remains at its original value $\frac{2}{3}$. In *principle*, successive applications of the $\Lambda$-operator should improve the solution, and, eventually, produce the exact solution. In fact, one can show that $\lim_{n\to\infty} \Lambda^n(1) = 0$ [see (**684**, 31)] so that an initial error $\varepsilon$ at any depth can ultimately be reduced to zero by repeated application of the $\Lambda$-operator. In practice, however, the convergence is too slow to be of value, for the effective range of the $\Lambda$-operator is of order $\Delta\tau = 1$, so errors at large depth are removed only "infinitely slowly." (We shall encounter this problem with $\Lambda$-iteration repeatedly in a wide variety of contexts! See, e.g.,



§6-1, §7-2.) Further, even a second application of the $\Lambda$-operator to $J_E^{(2)}(\tau)$ introduces the functions $\Lambda_\tau[E_n(t)]$, which are cumbersome to compute [see (**361**), equations (14-50), (14-53), and (37-36) through (37-44)]. Therefore alternative methods for obtaining a solution must be developed.

*Exercise 3-3*: Using the results of Table 3-2, evaluate the percentage errors of $J_E(\tau)$ and $J_E^{(2)}(\tau)$ and display them in a plot. The required values of $E_n(\tau)$ can be found in (**4**, 245).

*Exercise 3-4*: Show that, although $J_E(\tau)$ was derived assuming $F$ = constant, the flux computed from $J_E(\tau)$ via equation (2-59) is not constant; make a plot of the error $\Delta F(\tau)/F$.

*Exercise 3-5*: Apply the $X$-operator [cf. equations (2-62) and (2-65)] to $J_E(\tau)$ to obtain $K_E^{(2)}(\tau) = \frac{3}{16}F[\frac{4}{3}\tau + \frac{8}{9} - \frac{4}{3}E_4(\tau) + 2E_5(\tau)]$. Use this result to write an analytical expression for the variable Eddington factor $f(\tau) \equiv K(\tau)/J(\tau)$. Show that $f(\tau = \infty) = \frac{1}{3}$ and $f(\tau = 0) = \frac{17}{42} = 0.405$. Using the results of Table 3-2, evaluate the fractional error in $f(\tau)$ [recall equation (3-15)] and plot it.

*Exercise 3-6*: Show that the improved estimate of the emergent intensity obtained by using $J_E^{(2)}(\tau)$ is $I_E^{(2)}(0, \mu) = \frac{3}{4}F\{\frac{7}{12} + \frac{1}{2}\mu + (\frac{1}{3}\mu + \frac{1}{2}\mu^2)\ln[(1 + \mu)/\mu]\}$. Compare this result and $I_E(0, \mu)$ given by equation (3-36) with the exact result shown in Table 3-1, and plot a graph of their fractional errors.

### ITERATION: THE UNSÖLD PROCEDURE

The primary shortcoming of the $\Lambda$-iteration procedure is its failure to yield an improvement in the solution at great depth. Unsöld (**638**, 141) proposed an ingenious alternative method that overcomes this inadequacy and can be generalized to the nongrey case. The basic idea is to start from an initial estimate for the source function $B(\tau)$, and to derive a perturbation equation for a change $\Delta B(\tau)$ that more nearly satisfies the requirement of radiative equilibrium.

If we calculate the flux from the initial guess $B(\tau)$, we will find that it is a function of depth, $H(\tau)$, and not exactly constant unless $B(\tau)$ happened to be the exact solution of the problem. From the first-order moment equation (3-9) we then have

$$K(\tau) = \int_0^\tau H(\tau')\,d\tau' + C \tag{3-39}$$

If we make the Eddington approximation $J(\tau) = 3K(\tau)$ and evaluate $C$ by writing $J(0) = 2H(0)$ [cf. equation (3-34)] we obtain

$$J(\tau) \approx 3\int_0^\tau H(\tau')\,d\tau' + 2H(0) \tag{3-40}$$

But from the zero-order moment equation (3-8) we have

$$B(\tau) = J(\tau) - [dH(\tau)/d\tau] \tag{3-41}$$

so that

$$B(\tau) \approx 3 \int_0^\tau H(\tau')\, d\tau' + 2H(0) - [dH(\tau)/d\tau] \tag{3-42}$$

Equation (3-42) cannot be exact because of the approximations made in its derivation, but it can be used with sufficient accuracy to compute perturbations. In particular, suppose $\Delta B(\tau)$ is chosen just so the flux computed from $B(\tau) + \Delta B(\tau)$ *is* constant; thus

$$B(\tau) + \Delta B(\tau) \approx 3 \int_0^\tau H\, d\tau' + 2H \tag{3-43}$$

and by subtraction of equation (3-42) from (3-43) we obtain an expression for $\Delta B(\tau)$, namely

$$\Delta B(\tau) = 3 \int_0^\tau \Delta H(\tau')\, d\tau' + 2\Delta H(0) - [d\,\Delta H(\tau)/d\tau] \tag{3-44}$$

Thus if we know the flux errors $\Delta H(\tau) \equiv H - H(\tau)$ we can compute the correction $\Delta B(\tau)$; this correction is then applied and new values of the flux are computed, which give new (smaller!) errors $\Delta H$; the process is iterated until $H$ becomes constant and $\Delta B \to 0$ at all $\tau$. Equation (3-44) can be generalized for nongrey atmospheres: see equation (7-18). Unsöld's procedure is very powerful compared to $\Lambda$-iteration, for it provides a great improvement in the solution *at depth* as well as at the surface; this result is demonstrated in the following exercise.

*Exercise 3-7*: Assume a starting solution $B(\tau) = 3H(\tau + c)$; i.e., $q(\tau) \equiv c$. (a) Show that $\Delta H(\tau) \equiv H - H(\tau) = \frac{3}{2}H[E_4(\tau) - cE_3(\tau)]$. Obtain expressions for $\Delta H(0)$ and $d(\Delta H)/d\tau$. (b) Apply Unsöld's procedure and show that

$$\Delta B(\tau) = 3H\left[\frac{17}{24} - c - \frac{1}{2}cE_2(\tau) + \frac{1}{2}E_3(\tau) + \frac{3}{2}cE_4(\tau) - \frac{3}{2}E_5(\tau)\right]$$

(c) Show that, independent of the initial choice of $c$, the improved solution has $q(0) = \frac{7}{12} = 0.583$ (exact value 0.577) and $q(\infty) = \frac{17}{24} = 0.708$ (exact value = 0.710). (d) Show that, in contrast, the $\Lambda$-operator acting on $q \equiv c$ gives $q(0) = \frac{1}{2}c + \frac{1}{4}$, which agrees with Unsöld's value only if $c = \frac{2}{3}$, and $q(\infty) = c$, which shows no improvement whatever at depth.

### THE METHOD OF DISCRETE ORDINATES

The method to be described now furnishes a means of obtaining both approximate solutions and the exact solution of the grey problem. More important, it introduces the fundamental mathematical scheme that provides the basis for practically all modern techniques of solving transfer equations. Introducing the definition of $J$ [equation (1-4)] into equation (3-5), the transfer equation [illegible]d may be written in the form



$$\mu[\partial I(\tau, \mu)/\partial\tau] = I(\tau, \mu) - \frac{1}{2}\int_{-1}^{1} I(\tau, \mu)\, d\mu \tag{3-45}$$

which is classified as an *integro-differential equation*. The essential difficulty in obtaining the solution arises from the presence of the integral over angle. However, definite integrals such as that in equation (3-45) may be evaluated numerically by means of a *quadrature sum* consisting of the function evaluated at a finite set of points on the interval of integration times appropriate weights. Thus introducing $\{\mu_j\}$ on $[-1, 1]$ we write, for any function $f(\mu)$,

$$\frac{1}{2}\int_{-1}^{1} f(\mu)\, d\mu \approx \frac{1}{2}\sum_{j=-n}^{n} a_j f(\mu_j) \tag{3-46}$$

The numbers $\{\mu_j\}$ are called the quadrature *points*, $\{a_j\}$ the quadrature *weights*, and $\{f(\mu_j)\}$ the *discrete ordinates*. Having chosen a definite quadrature formula, we replace the integro-differential transfer equation (3-45) with a coupled system of $2n$ ordinary differential equations:

$$\mu_i(\partial I_i/\partial\tau) = I_i - \frac{1}{2}\sum_{j=-n}^{n} a_j I_j, \quad (i = \pm 1, \ldots, \pm n) \tag{3-47}$$

where $I_i$ denotes $I(\tau, \mu_i)$. The radiation field is no longer represented as a *continuous* function, but rather in terms of a set of *pencils* of radiation, each of which represents the value of $I(\mu)$ over a definite interval. On physical grounds it is reasonable to expect the solution to become exact in the limit $n \to \infty$.

The accuracy of the quadrature depends both upon the number of points, and upon their distribution on the interval. If the points are distributed uniformly on the interval we obtain a *Newton–Cotes* formula, of which *Simpson's rule* with points at $\{\mu_i\} = (-1, 0, 1)$ is a familiar example. A better choice is to use a *Gaussian* formula, in which the $2n$ points on $[-1, 1]$ are chosen to be the roots of the Legendre polynomial of order $2n$. It would take us too far afield to discuss the construction and accuracy of quadrature formulae [see (**161**, Chap. 2)]; an important result that we shall merely state is this: an $m$-point Newton–Cotes formula gives exact results for polynomials of order $m - 1$ (for $m$ even) or $m$ (for $m$ odd), but an $m$-point Gauss formula is exact for polynomials of order $2m - 1$. For solving the transfer equation the *double-Gauss* formula is superior (**619**) to the ordinary (or "single") Gauss formula. Here one chooses two *separate* $n$-point quadratures on the ranges $(-1 \leqslant \mu \leqslant 0)$ and $(0 \leqslant \mu \leqslant 1)$; on each range the $n$ points are given by the roots of the Legendre polynomial of order $n$, scaled from $[-1, 1]$ to

the appropriate range. This approach has the advantage that $I(\tau, +\mu)$ and $I(\tau, -\mu)$ are approximated independently, and thus the integration formula can account, without difficulty, for the physical fact that $I(-\mu) \equiv 0$ at $\tau = 0$ while $I(+\mu)$ remains finite. In the single-Gauss formula, the discontinuity in $I(\mu)$ at $\mu = 0$ when $\tau = 0$ introduces significant errors. In all of these formulae the points are chosen symmetrically about zero so that $\mu_{-j} = -\mu_j$, while $a_{-j} = a_j$.

We now wish to solve the system of equations (3-47). Observing that the system is linear and of the first order, we take a trial solution of the form $I_i = g_i \exp(-k\tau)$, where $g_i$ and $k$ are to be specified. Substituting into equation (3-47) we find

$$g_i(1 + k\mu_i) = \frac{1}{2} \sum_{j=-n}^{n} a_j g_j = c \tag{3-48}$$

so that $g_i = c/(1 + k\mu_i)$. If we use this specific form for $g_i$ and again substitute the trial solution into equation (3-47), we find

$$\frac{1}{2} \sum_{j=-n}^{n} a_j (1 + k\mu_j)^{-1} = 1 \tag{3-49}$$

This is the *characteristic equation*, which is satisfied only by certain values of $k$, called the *characteristic roots* (eigenvalues). Recalling that $a_{-j} = a_j$ and $\mu_{-j} = -\mu_j$, equation (3-49) can be used to define the *characteristic function* $T(k^2)$,

$$T(k^2) \equiv 1 - \sum_{j=1}^{n} a_j (1 - k^2 \mu_j^2)^{-1} \tag{3-50}$$

The roots of $T$ [i.e., those values of $k$ for which $T(k^2) = 0$] are the desired characteristic roots. If we set $f(\mu) \equiv 1$ in equation (3-46) we see that

$$\frac{1}{2} \sum_{j=-n}^{n} a_j = \sum_{j=1}^{n} a_j = 1$$

hence it follows that $k^2 = 0$ is a solution of the characteristic equation; i.e., $T(k^2 = 0) = 0$. There are an additional $(n - 1)$ nonzero roots, which may be seen as follows. Note that $k^2 = \mu_j^{-2}$ is a *pole* of $T$, which becomes infinite for these values of $k^2$. For $k^2 = \mu_j^{-2} - \varepsilon$, $T(k^2) < 0$, and by making $\varepsilon$ arbitrarily small, $T(k^2) \to -\infty$. Similarly, for $k^2 = \mu_j^{-2} + \varepsilon$, $T(k^2) > 0$, and as $\varepsilon \to 0$, $T \to +\infty$. It is thus clear that $T$ must pass through zero somewhere on the interval between successive poles, hence the $(n - 1)$ nonzero roots must satisfy the relations

$$\mu_1^{-2} < k_1^2 < \mu_2^{-2} < \cdots < k_{n-1}^2 < \mu_n^{-2}$$



where we have ordered $\{\mu_i\}$ such that $\mu_i > \mu_{i+1}$. Note that the largest $\mu_i$ must be less than unity, hence the smallest nonzero $k^2$ must be greater than unity. In all there exist $2n - 2$ nonzero values of the $k$'s, in pairs of the form $\pm k_i$ $(i = 1, \ldots, n - 1)$.

The *general* solution of the system (3-47) is therefore of the form

$$I_i(\tau) = b\left[\sum_{\alpha=1}^{n-1} L_\alpha (1 + k_\alpha \mu_i)^{-1} e^{-k_\alpha \tau} + \sum_{\alpha=1}^{n-1} L_{-\alpha} (1 - k_\alpha \mu_i)^{-1} e^{+k_\alpha \tau}\right] \tag{3-51}$$

We must now seek the special solution corresponding to the root $k^2 = 0$. In view of equation (3-11) which shows that $J(\tau)$ must become linear in $\tau$ at depth, we examine trial solutions of the form $I_i = b(\tau + q_i)$. Substituting this expression into equation (3-47) we obtain

$$q_i = \mu_i + \frac{1}{2} \sum_{j=-n}^{n} a_j q_j \tag{3-52}$$

Now observing that if we set $f(\mu) = \mu$ in equation (3-46) the quadrature sum $\sum a_j \mu_j$ is zero, we see that equation (3-52) is satisfied by the simple choice $q_i = Q + \mu_i$. Hence the particular solution is $I_i(\tau) = b(\tau + Q + \mu_i)$ and the *complete* solution is

$$I_i(\tau) = b\left(\tau + Q + \mu_i + \sum_{\alpha=1}^{n-1} \frac{L_\alpha e^{-k_\alpha \tau}}{1 + k_\alpha \mu_i} + \sum_{\alpha=1}^{n-1} \frac{L_{-\alpha} e^{+k_\alpha \tau}}{1 - k_\alpha \mu_i}\right) \tag{3-53}$$

We must now specify the $2n$ unknown coefficients $b$, $Q$, and $L_{\pm\alpha}$. This is done by application of the boundary conditions.

In the case of a *finite* slab of total optical thickness $T$, both $I_{-i}^{-} \equiv I(0, -\mu_i)$ and $I_i^{+} \equiv I(T, +\mu_i)$ are *given* functions of $\mu$. Thus we may write $2n$ equations for the $2n$ unknowns:

$$I(0, -\mu_i) = b\left(Q - \mu_i + \sum_{\alpha=1}^{n-1} \frac{L_\alpha}{1 - k_\alpha \mu_i} + \sum_{\alpha=1}^{n-1} \frac{M_\alpha e^{-k_\alpha T}}{1 + k_\alpha \mu_i}\right) = I_{-i}^{-} \tag{3-54a}$$

and

$$I(T, +\mu_i) = b\left(Q + T + \mu_i + \sum_{\alpha=1}^{n-1} \frac{L_\alpha e^{-k_\alpha T}}{1 + k_\alpha \mu_i} + \sum_{\alpha=1}^{n-1} \frac{M_\alpha}{1 - k_\alpha \mu_i}\right) = I_i^{+} \tag{3-54b}$$

where, to improve the numerical condition of the equations, $M_\alpha \equiv L_{-\alpha} e^{k_\alpha T}$. Equations (3-54) may be solved by standard numerical techniques for linear equations.

In the case of a *semi-infinite* atmosphere in radiative equilibrium we have the boundary conditions $I_{-i}^{-} = I(0, -\mu_i) \equiv 0$, and demand that $I(\tau)$ must not diverge exponentially as $\tau \to \infty$. To satisfy the latter constraint we set

$$L_{-\alpha} \equiv 0 \tag{3-55}$$

and use the upper boundary condition to write the $n$ equations

$$Q - \mu_i + \sum_{\alpha=1}^{n-1} L_\alpha (1 - k_\alpha \mu_i)^{-1} = 0, \quad (i = 1, \ldots, n) \tag{3-56}$$

Solution of equations (3-56) yields the $n$ unknowns $Q$ and $L_\alpha$. In addition, we require that the flux equal the nominal flux $F$. Thus we demand that

$$2 \int_{-1}^{1} I(\mu) \mu \, d\mu = 2 \sum_{j=-n}^{n} a_j \mu_j I_j = F \tag{3-57}$$

Substituting equations (3-53) (with $L_{-\alpha} \equiv 0$), we have

$$2b \left[ (\tau + Q) \sum_{j=-n}^{n} a_j \mu_j + \sum_{j=-n}^{n} a_j \mu_j^2 + \sum_{\alpha=1}^{n-1} L_\alpha e^{-k_\alpha \tau} \sum_{j=-n}^{n} \frac{a_j \mu_j}{1 + k_\alpha \mu_j} \right] = F \tag{3-58}$$

In view of equation (3-46) the first sum is zero, the second equals $\frac{2}{3}$, and from the characteristic equation the fourth is found to be zero.

*Exercise 3-8*: Verify the statement made above about the values of the sums in equation (3-58).

Thus we find that $b = \frac{3}{4}F$, as would be expected from equation (3-11); note also that the quadrature calculation yields a *constant* flux automatically. Finally, the complete solution for the semi-infinite atmosphere may be written

$$I_i(\tau) = \frac{3}{4} F \left[ \tau + Q + \mu_i + \sum_{\alpha=1}^{n-1} L_\alpha e^{-k_\alpha \tau} (1 + k_\alpha \mu_i)^{-1} \right], \quad (i = 1, \ldots, n) \tag{3-59}$$

We may compute $J(\tau)$ by substituting equations (3-59) into the quadrature formula; making use of the characteristic equation (3-49), we obtain

$$J(\tau) = \frac{3}{4} F \left( \tau + Q + \sum_{\alpha=1}^{n-1} L_\alpha e^{-k_\alpha \tau} \right) \tag{3-60}$$



and thus the discrete-ordinate representation of $q(\tau)$ is

$$q(\tau) = Q + \sum_{\alpha=1}^{n-1} L_\alpha e^{-k_\alpha \tau} \tag{3-61}$$

*Exercise 3-9*: Derive equation (3-60).

Numerical results for $q(\tau)$ were obtained for $n = 1, \ldots, 4$ by Chandrasekhar (**153**) using a single-Gauss formula and by Sykes (**619**) using the superior double-Gauss formula. In every case the *exact* value $q(0) = 1/\sqrt{3}$ is obtained. The maximum percentage (absolute) error in $J(\tau)$ for the single-Gauss solutions is (9.0, 4.1, 2.7, 2.0) for $n = 1, \ldots, 4$, respectively, while for the double-Gauss it is (9.0, 1.8, 0.9) for $n = 1, \ldots, 3$, respectively. The double-Gauss estimates of $q(\infty)$, namely $Q$, are 0.71132 and 0.71057 for $n = 2$ and 3; these values compare favorably with the exact value 0.710446. The double-Gauss solution with $n = 3$ gives the emergent intensity $I(0, \mu)/F$ with a root-mean-square error of only 0.1%, and is very accurate (0.02%) for $\mu \geqslant 0.3$. It should be stressed that the main importance of the discrete ordinate method is that in the limit $n \to \infty$ it yields the exact solution and that it affords an extremely powerful approximation technique for more complicated problems.

Several very important results that will be useful in later work can be established by analytical manipulation of the characteristic equation and the boundary conditions. To simplify the notation we define $x \equiv 1/k$ and $X \equiv 1/k^2$. The characteristic function defined in equation (3-50) may be rewritten in the equivalent forms

$$T(X) \equiv 1 - X \sum_{j=1}^{n} \frac{a_j}{X - \mu_j^2} = \sum_{j=1}^{n} a_j - X \sum_{j=1}^{n} \frac{a_j}{X - \mu_j^2} = \sum_{j=1}^{n} \frac{a_j \mu_j^2}{\mu_j^2 - X} \tag{3-62}$$

To clear $T(X)$ of fractions, multiply through by $\prod_{j=1}^{n} (\mu_j^2 - X)$; this yields

$$P(X) \equiv \prod_{j=1}^{n} (\mu_j^2 - X) T(X) = \sum_{i=1}^{n} a_i \mu_i^2 \prod_{j \neq i}^{n} (\mu_j^2 - X) \tag{3-63}$$

which is clearly a polynomial of order $(n - 1)$ in $X$. But we know that $T(X)$ has the $(n - 1)$ roots $X_1 = 1/k_1^2, \ldots, X_{n-1} = 1/k_{n-1}^2$, so $P(X)$ must have the form $C(X - X_1) \cdots (X - X_{n-1})$. To evaluate the constant, we note that the coefficient of the term in $X^{n-1}$ in equation (3-63) is $(-1)^{n-1} \sum_{i=1}^{n} a_i \mu_i^2 = (-1)^{n-1} \frac{1}{3}$; this is simply $C$ itself. We thus have

$$P(X) = \frac{1}{3} (X_1 - X) \cdots (X_{n-1} - X)$$

and therefore

$$T(X) = \frac{1}{3}\left[\prod_{j=1}^{n-1}(X_j - X)\right] \Big/ \prod_{j=1}^{n}(\mu_j^2 - X) \tag{3-64}$$

From equation (3-62) we see that $T(X = 0) = 1$, hence setting $X = 0$ in equation (3-64) we obtain the useful (see below) result that

$$\mu_1\mu_2 \cdots \mu_n k_1 k_2 \cdots k_{n-1} = 1/\sqrt{3} \tag{3-65}$$

Now consider the emergent intensity $I(0, \mu)$. Define a function $S(\mu)$ such that

$$S(\mu_i) \equiv Q - \mu_i + \sum_{\alpha=1}^{n-1} L_\alpha (1 - k_\alpha \mu_i)^{-1} \tag{3-66}$$

The surface boundary conditions in equation (3-56) may then be written

$$I(0, -\mu_i) = \frac{3}{4} F S(\mu_i) = 0 \tag{3-67}$$

We then generalize $S(\mu)$ to apply at all values of $\mu$ and write

$$I(0, \mu) = \frac{3}{4} F S(-\mu) \tag{3-68}$$

for $\mu \geqslant 0$. Note that, with this generalization, $I(0, -\mu)$ is not $\equiv 0$, but will in general have nonzero values for $-\mu \neq \mu_{-i}$. By working with $S(\mu)$, we can obtain an expression for $I(0, \mu)$ that does not involve the constants $L_\alpha$ and $Q$ explicitly. Clear fractions from equation (3-66) to obtain

$$S(\mu)\prod_{\alpha=1}^{n-1}(1 - k_\alpha \mu) = (Q - \mu)\prod_{\alpha=1}^{n-1}(1 - k_\alpha \mu) + \sum_{\alpha=1}^{n-1} L_\alpha \prod_{i\ \alpha}^{n-1}(1 - k_i \mu) \tag{3-69}$$

The righthand expression is obviously a polynomial of order $n$ in $\mu$. But $S(\mu)$ has the $n$ roots $\mu_1, \ldots, \mu_n$; hence this polynomial must be of the form $C(\mu - \mu_1) \cdots (\mu - \mu_n)$. To find $C$ we note that the coefficient $\mu^n$ on the righthand side of equation (3-69) is $(-1)^n k_1 \cdots k_{n-1}$, which is $C$ itself. Therefore

$$S(\mu) = \frac{(\prod_{i=1}^{n-1} k_i)\prod_{i=1}^{n}(\mu_i - \mu)}{\prod_{i=1}^{n-1}(1 - k_i \mu)} = \frac{\prod_{i=1}^{n}(\mu_i - \mu)}{\prod_{i=1}^{n-1}(x_i - \mu)} \tag{3-70}$$

which, when inserted into equation (3-68), yields the desired expression for $I(0, \mu)$. It is customary to define a *limb-darkening function* $H(\mu)$ as

$$H(\mu) \equiv I(0, \mu)/I(0, 0) \tag{3-71}$$



[note that unlike equation (3-37), the reference point here is the limb, not disk-center]. In the discrete-ordinate approximation we find from equations (3-68) and (3-70)

$$H(\mu) = \left[\prod_{i=1}^{n}(1 + \mu_i^{-1}\mu)\right] \Big/ \prod_{i=1}^{n-1}(1 + k_i \mu) \tag{3-72}$$

By further analysis using $S(\mu)$ it is possible to write explicit expressions for the $L_\alpha$'s and $Q$ in terms of the points $\{\mu_i\}$ and the roots $\{k_\alpha\}$ [see (**161**, 78–79; **684**, 25)].

Before leaving the discrete ordinate method, let us show that $q(0) = 1/\sqrt{3}$ is the *exact* value. First, note that in the $n$th discrete approximation

$$J_n(0) = \frac{3}{4} F\left(Q + \sum_{\alpha=1}^{n-1} L_\alpha\right) = \frac{3}{4} F q(0) \tag{3-73}$$

while from equation (3-59) we find

$$I_n(0, 0) = \frac{3}{4} F\left(Q + \sum_{\alpha=1}^{n-1} L_\alpha\right) = J_n(0) \tag{3-74}$$

*independent of the order n*. Thus we conclude that in the *exact* solution $J(0) = I(0, 0)$. But from equations (3-68) and (3-70)

$$I_n(0, 0) = \frac{3}{4} F S(0) = \frac{3}{4} F \mu_1 \cdots \mu_n k_1 \cdots k_{n-1} \tag{3-75}$$

Then, combining equations (3-65), (3-75), and (3-73), we deduce that, *independent of the order n*, $q_n(0) = 1/\sqrt{3}$; hence this result is exact.

## 3-4 Exact Solution

The exact solution for $q(\tau)$ and $H(\mu)$ can be obtained by taking the discrete-ordinate method to the limit $n \to \infty$ (**161**, Chap. 5; **361**, §27), by applying the *principle of invariance* (**161**, Chap. 4; **361**, §28), or by direct *Laplace transform* methods (**361**, §29; **684**, Chap 3). Only certain important results will be quoted here, and the reader should consult the references cited for details.

Several expressions for $H(\mu)$ exist (**361**, 186–187); a form convenient for numerical computation is

$$H(\mu) = (\mu + 1)^{-1/2} \exp\left[\frac{1}{\pi}\int_0^{\pi/2} \frac{\theta \tan^{-1}(\mu \tan \theta)}{1 - \theta \cot \theta}\, d\theta\right] \tag{3-76}$$

TABLE 3-1
*Exact Limb-Darkening Law for Grey Atmospheres*

| $\mu$ | $I(0, \mu)/F$ | $\mu$ | $I(0, \mu)/F$ |
|---|---|---|---|
| 0.0 | 0.43301 | 0.6 | 0.95009 |
| 0.1 | 0.54012 | 0.7 | 1.02796 |
| 0.2 | 0.62802 | 0.8 | 1.10535 |
| 0.3 | 0.71123 | 0.9 | 1.18238 |
| 0.4 | 0.79210 | 1.0 | 1.25912 |
| 0.5 | 0.87156 | | |

Evaluation of this integral (**152**; **518**) yields the results summarized in Table 3-1.

The value of $q(\infty)$ can be obtained by noting from equation (3-15) that $K(0) = Hq(\infty)$, and thus, using equation (3-71),

$$q(\infty) = \left[\int_0^1 H(\mu)\mu^2\, d\mu\right] \Big/ \int_0^1 H(\mu)\mu\, d\mu \tag{3-77}$$

From the known expressions for $H(\mu)$ one then obtains

$$q(\infty) = \frac{6}{\pi^2} + \frac{1}{\pi}\int_0^{\pi/2}\left(\frac{3}{\theta^2} - \frac{1}{1 - \theta\cot\theta}\right) d\theta \tag{3-78}$$

which yields $q(\infty) = 0.71044609$ (**519**). Finally, a closed-form expression can

TABLE 3-2
*The Exact Solution for $q(\tau)$*

| $\tau$ | $q(\tau)$ | $\tau$ | $q(\tau)$ |
|---|---|---|---|
| 0.00 | 0.577351 | 0.8 | 0.693534 |
| 0.01 | 0.588236 | 1.0 | 0.698540 |
| 0.03 | 0.601242 | 1.5 | 0.705130 |
| 0.05 | 0.610758 | 2.0 | 0.707916 |
| 0.10 | 0.627919 | 2.5 | 0.709191 |
| 0.20 | 0.649550 | 3.0 | 0.709806 |
| 0.30 | 0.663365 | 3.5 | 0.710120 |
| 0.40 | 0.673090 | 4.0 | 0.710270 |
| 0.50 | 0.680240 | 5.0 | 0.710398 |
| 0.60 | 0.685801 | $\infty$ | 0.710446 |



be written (**407**) for $q(\tau)$, namely

$$q(\tau) = q(\infty) - \frac{1}{2\sqrt{3}}\int_0^1 \frac{e^{-\tau/u}\, du}{H(u)Z(u)} \tag{3-79}$$

where $H(u)$ is as defined above and

$$Z(u) \equiv \left[1 - \frac{1}{2}u\ln\left(\frac{1+u}{1-u}\right)\right]^2 + \frac{1}{4}\pi^2u^2 \tag{3-80}$$

Results obtained from a numerical evaluation of equation (3-79) are given in Table 3-2.

## 3-5 Emergent Flux from a Grey Atmosphere

The basic physical assumption made in the grey-atmosphere problem is that the opacity is independent of frequency. In this event, the constraint of radiative equilibrium reduces to $S(\tau) \equiv J(\tau)$, and the problem simplifies to that of obtaining the solution of equation (3-6). If, in addition, it is assumed that LTE prevails, then we may equate $B(\tau) = (\sigma_R T^4)/\pi$ to $J(\tau)$, and thus arrive at equation (3-16) for $T(\tau)$. The radiation field has a dependence upon frequency because the source function, which we assume is $B_\nu(\tau)$, depends upon frequency. Given the source function, the flux, also frequency-dependent, can be computed at any depth by means of equation (2-59), which now reads

$$F_\nu(\tau) = 2\int_\tau^\infty B_\nu[T(t)]E_2(t - \tau)\, dt - 2\int_0^\tau B_\nu[T(t)]E_2(\tau - t)\, dt \tag{3-81}$$

The temperature appears in the Planck function only in the combination $(h\nu/kT)$; further, the ratio $T(\tau)/T_{\text{eff}}$ is a unique function of $\tau$ [cf. equation (3-16)]—say, $1/p(\tau)$. We may therefore simplify the equation by introducing a parameter $\alpha \equiv (h\nu/kT_{\text{eff}})$, in terms of which we can write $(h\nu/kT) = \alpha p(\tau)$. Expressing the flux in the same units by writing $F_\alpha(\tau) \equiv F_\nu(\tau)(d\nu/d\alpha)$, and using the relation $F \equiv (\sigma_R T_{\text{eff}}^4)/\pi$, we may rewrite equation (3-81) as

$$\frac{F_\alpha(\tau)}{F} = \left(\frac{4\pi k^4}{h^3c^2\sigma_R}\right)\alpha^3\left\{\int_\tau^\infty \frac{E_2(t - \tau)\, dt}{\exp[\alpha p(t)] - 1} - \int_0^\tau \frac{E_2(\tau - t)\, dt}{\exp[\alpha p(t)] - 1}\right\} \tag{3-82}$$

The expression in the brackets is a function of $\alpha$ and $\tau$ only, and may be calculated once and for all. A tabulation of $F_\alpha(\tau)/F$ is given in (**161**, 295), and a plot of the function is displayed in Figure 3-1. The figure shows clearly the degradation of photon energies as they transfer from depth to the surface; for example, the most common photon energy at $\tau = 0$ is only about

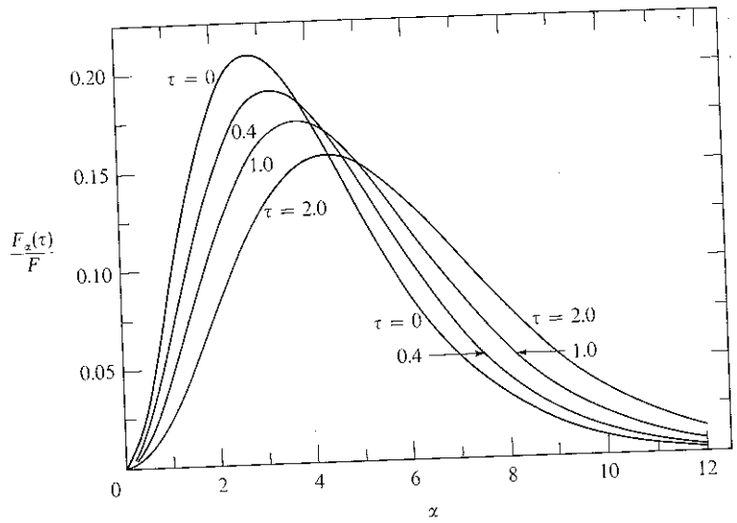



FIGURE 3-1
Frequency distribution of flux at selected depths in a grey atmosphere. From (**155**), by permission.

75 percent of that at $\tau = 1$. This progressive reddening of photons in the outer layers results from the outward decrease in temperature produced by the requirement of radiative equilibrium.

## 3-6 Small Departures from Greyness

By use of an appropriate mean opacity, it is possible to account for small departures from greyness, at least approximately, and thus to extend greatly the usefulness of the results obtained for a grey atmosphere. Suppose that the frequency variation of the opacity is the same at all depths so that we can write

$$\chi_\nu = \overline{\chi}_c(1 + \beta_\nu) \equiv \overline{\chi}_c\gamma_\nu \tag{3-83}$$

where

$$\overline{\chi}_c F \equiv \int_0^\infty \chi_\nu F_\nu^{(1)}\, d\nu \tag{3-84}$$

In equation (3-84), $F_\nu^{(1)}$ denotes *the flux in a grey atmosphere*. The mean opacity $\overline{\chi}_c$ defined in equation (3-84) is known as the *Chandrasekhar mean*; as we shall see in what follows, this mean is constructed in a way that makes optimum use of the information contained in the grey solution. Unlike the flux mean [equation (3-21)] the Chandrasekhar mean can be computed straightaway for any given frequency dependence (i.e., $\beta_\nu$ or $\gamma_\nu$) of the opacity because $F_\nu^{(1)}$ is a *known* function.



Let us now consider how we may solve the nongrey transfer problem using a method of successive approximations. The nongrey transfer equation (assuming LTE) is

$$\mu(\partial I_\nu/\partial\overline{\tau}) = (\chi_\nu/\overline{\chi})(I_\nu - B_\nu) = (1 + \beta_\nu)(I_\nu - B_\nu) \tag{3-85}$$

If we suppose that the departures $\beta_\nu$ may be regarded as small, then a first approximation to the solution of equation (3-85) is obtained by setting $\beta_\nu \equiv 0$ initially. The transfer equation then returns to the equation for the grey problem itself, namely

$$\mu(\partial I_\nu^{(1)}/\partial\overline{\tau}) = I_\nu^{(1)} - B_\nu \tag{3-86}$$

whose solution is already known. To obtain a second approximation we write

$$\begin{aligned}\mu(\partial I_\nu^{(2)}/\partial\overline{\tau}) &= I_\nu^{(2)} - B_\nu + \beta_\nu(I_\nu^{(1)} - B_\nu) \\ &= I_\nu^{(2)} - B_\nu + \beta_\nu\mu(\partial I_\nu^{(1)}/\partial\overline{\tau})\end{aligned} \tag{3-87}$$

by substitution from equation (3-86). If we demand that the radiation field resulting from this second approximation should satisfy the constraint of radiative equilibrium, then we must have $(dF^{(2)}/d\overline{\tau}) \equiv 0$, where $F^{(2)}$ is the integrated flux; then from equation (3-87)

$$J^{(2)} - B + \frac{d}{d\overline{\tau}}\left[\int_0^\infty \beta_\nu F_\nu^{(1)}\, d\nu\right] = 0 \tag{3-88}$$

But note that equations (3-83) and (3-84) imply that

$$\overline{\chi}_c F = \overline{\chi}_c \int_0^\infty (1 + \beta_\nu)F_\nu^{(1)}\, d\nu = \overline{\chi}_c F + \overline{\chi}_c \int_0^\infty \beta_\nu F_\nu^{(1)}\, d\nu \tag{3-89}$$

or that

$$\int_0^\infty \beta_\nu F_\nu^{(1)}\, d\nu \equiv 0 \tag{3-90}$$

Therefore the radiative equilibrium constraint for a nongrey atmosphere treated by the above approximation scheme collapses to $J^{(2)} = B(\overline{\tau})$. This shows that the *grey* solution for $T(\overline{\tau})$ on the Chandrasekhar-mean optical-depth scale will *automatically* satisfy the condition of radiative equilibrium in the first approximation to the nongrey atmosphere. The method for obtaining higher approximations is described in (**161**, 296 ff.)

At this first level of approximation we may compute the nongrey emergent intensity as

$$I_\nu(0, \mu) = \int_0^\infty B_\nu[T(\overline{\tau})] \exp(-\gamma_\nu\overline{\tau}/\mu)(\gamma_\nu/\mu)\, d\overline{\tau} \tag{3-91}$$

and the flux as

$$F_\nu(0) = 2\int_0^\infty B_\nu[T(\bar{\tau})]E_2(\gamma_\nu\bar{\tau})\gamma_\nu\, d\bar{\tau} \tag{3-92}$$

If we introduce the parameter $\alpha \equiv (h\nu/kT_{\text{eff}})$ as was done in §3-5, and write

$$B_\nu[T(\bar{\tau})] = B_\nu(T_0)\left[\frac{\exp(\alpha T_{\text{eff}}/T_0) - 1}{\exp[\alpha p(\bar{\tau})] - 1}\right] = B_\nu(T_0)b_\alpha(\bar{\tau}) \tag{3-93}$$

equations (3-91) and (3-92) reduce to the parametric forms

$$I_\nu(0, \mu) = B_\nu(T_0)\int_0^\infty b_\alpha(\bar{\tau})\exp(-\gamma_\nu\bar{\tau}/\mu)(\gamma_\nu/\mu)\, d\bar{\tau} \equiv B_\nu(T_0)\mathscr{I}(\alpha, \gamma_\nu/\mu) \tag{3-94}$$

and

$$F_\nu(0) = 2B_\nu(T_0)\int_0^\infty b_\alpha(\bar{\tau})E_2(\gamma_\nu\bar{\tau})\gamma_\nu\, d\bar{\tau} \equiv B_\nu(T_0)\mathscr{F}(\alpha, \gamma_\nu) \tag{3-95}$$

The function $b_\alpha(\bar{\tau})$ is unique for a given value of $\alpha$, and hence the functions $\mathscr{I}(\alpha, \beta)$ and $\mathscr{F}(\alpha, \beta)$ may be computed once and for all; tables of these functions are given in (**161**, 306–307).

The functions $\mathscr{I}(\alpha, \beta)$ and $\mathscr{F}(\alpha, \beta)$ described above at one point played an important role in the development of the theory of stellar atmospheres. By analysing the observational material available for the sun, Münch (**473**) was able to find those values of $\gamma_\nu$ that best reproduced the observed fluxes and limb-darkening. These were shown to be compatible with the frequency variation of the absorption coefficient of the negative hydrogen ion, $H^-$, as computed by Chandrasekhar. Analyses such as these led to conclusive identification of $H^-$ as the major opacity source in the solar atmosphere. We examine opacities in some detail in the next chapter.

# 4

# *Absorption Cross-Sections*

In this chapter we outline the quantum mechanical calculation of atomic absorption cross-sections. The discussion is meant to be self-contained, but limitations of space require that knowledge of the basic principles of quantum mechanics at the level of (**392**, Chaps. 2–8) be presupposed.

## 4-1 The Einstein Relations for Bound–Bound Transitions

Let us first consider the absorption and emission of radiation by an atom in a transition between two bound states. Assume that the lower state ($i$) has statistical weight $g_i$, and the upper state ($j$) statistical weight $g_j$. There are three basic processes involved, which are usually described in terms of rate coefficients first introduced by Einstein (**207**).

The first process is the direct *absorption* of radiation, leading to an upward transition from level $i$ to level $j$. The rate at which this process occurs for radiation of specific intensity $I_\nu$ can be written in terms of the *Einstein coefficient* $B_{ij}$ as

$$n_i(\nu)R_{ij}(d\omega/4\pi) = n_i(\nu)B_{ij}I_\nu(d\omega/4\pi) \tag{4-1}$$

where $n_i(v)$ is the number of atoms per cm$^3$ in state $i$ that can absorb radiation at frequencies on the range $(v, v + dv)$. In general the spectrum line corresponding to the transition will not be sharp; rather, because of perturbations exerted by nearby atoms and ions, as well as the finite lifetime of the upper state, it will have a spread in frequency that we can describe by an *absorption profile*, $\phi_v$, normalized such that $\int \phi_v \, dv = 1$. Thus if the total number of atoms in state $i$ is $n_i$, the number capable of absorbing at frequency $v$ is $n_i(v) = n_i\phi_v$. In making the transition from level $i$ to level $j$, the atom absorbs photons of energy $hv_{ij} = E_j - E_i$. Thus the rate at which *energy* is removed from an incident beam of radiation is

$$a_v I_v = n_i(B_{ij}hv_{ij}/4\pi)\phi_v I_v \tag{4-2}$$

where $a_v$ denotes a macroscopic absorption coefficient (per unit volume), uncorrected for stimulated emission (see below).

For atoms returning from level $j$ to level $i$, two processes are possible. The first of these is a spontaneous transition with the emission of a photon. Writing the *probability of spontaneous emission* per unit time as $A_{ji}$, the rate of emission of energy is

$$\eta_v(\text{spontaneous}) = n_j(A_{ji}hv_{ij}/4\pi)\psi_v \tag{4-3}$$

Here the emission profile $\psi_v$ specifies the number of atoms in the upper state that can emit photons on the frequency range $(v, v + dv)$; it is normalized such that $\int \psi_v \, dv = 1$. The other possible return process is a transition *induced by the radiation field* (*stimulated emission*). The rate at which such emission occurs is assumed proportional to the intensity of the incident radiation field. The energy emitted may be written in terms of the Einstein coefficient $B_{ji}$ as

$$\eta_v(\text{induced}) = n_j(B_{ji}hv_{ij}/4\pi)\psi_v I_v \tag{4-4}$$

In writing equation (4-4), use has been made of the result that the profile for induced emission is the same as that for spontaneous emission, as can be shown from general quantum mechanical considerations (**197**, §62). It should be noted that spontaneous emission takes place *isotropically*. Induced emission is proportional to and has the same angular distribution as $I_v$; because of this, induced emissions are sometimes considered to be *negative absorptions*, though this is not quite correct, for in general $\psi_v$ will not be identical to $\phi_v$.

The coefficients $A_{ji}$, $B_{ji}$, and $B_{ij}$ are simply related, as can be shown by calculating rates of absorption and emission in thermodynamic equilibrium. In strict T.E., the radiation field is isotropic, and $I_v \equiv B_v$, the Planck function. Furthermore, in T.E. the occupation numbers of levels $i$ and $j$ are related



by the Boltzmann law [cf. equation (5-5)]

$$(n_j/n_i)^* = (g_j/g_i)\exp(-hv_{ij}/kT) \tag{4-5}$$

Moreover, in T.E., $\psi_v = \psi_v^* = \phi_v$ [cf. equation (2-14)]. Now in strict T.E., each upward transition on the range $(v, v + dv)$ must be exactly balanced (detailed balancing) by an emission on that range. Hence, frequency by frequency we must have

$$n_i^* B_{ij} B_v = n_j^* A_{ji} + n_j^* B_{ji} B_v \tag{4-6}$$

Solving for $B_v$, we find

$$B_v = \frac{n_j^* A_{ji}}{n_i^* B_{ij} - n_j^* B_{ji}} = \left(\frac{A_{ji}}{B_{ji}}\right)\left[\left(\frac{g_i B_{ij}}{g_j B_{ji}}\right)\exp(hv_{ij}/kT) - 1\right]^{-1} \tag{4-7}$$

But the correct expression for the Planck function as obtained from statistical mechanical arguments is $B_v = (2hv^3/c^2)[\exp(hv/kT) - 1]^{-1}$, and to make this expression correspond to equation (4-7) we conclude that

$$A_{ji} \equiv (2hv^3/c^2)B_{ji} \tag{4-8}$$

and

$$g_i B_{ij} \equiv g_j B_{ji} \tag{4-9}$$

We shall use these results repeatedly in later work.

It should be noted that, although the *Einstein relations* in equations (4-8) and (4-9) were derived, for ease, from a thermodynamic equilibrium argument, the Einstein coefficients are really *properties of the atom only*, and must be independent of the nature of the radiation field. Therefore we must conclude that equations (4-8) and (4-9) are true in general. It is of interest that historically a line of argument similar to that presented above led to the realization that the stimulated emission process must occur in nature—a fact not intuitively obvious at first sight.

*Exercise 4-1*: Show that for a Planckian radiation field at typical stellar temperatures ($\sim 10^4$ °K) spontaneous emissions occur much more rapidly than induced emissions in the ultraviolet where $hv/kT \gg 1$, while the reverse is true at far-infrared and radio wavelengths where $hv/kT \ll 1$.

The microscopic formulation described above may immediately be incorporated into the equation of transfer. If *only* the bound–bound process occurs, then the appropriate transfer equation is

$$\mu(\partial I_v/\partial z) = [n_j A_{ji}\psi_v - (n_i B_{ij}\phi_v - n_j B_{ji}\psi_v)I_v](hv_{ij}/4\pi) \tag{4-10}$$

Here we have followed the usual practice of grouping together all terms involving $I_\nu$. In this way one can define a *line absorption coefficient corrected for stimulated emission*, namely

$$\chi_l(\nu) = \left(\frac{n_i B_{ij} h\nu_{ij}}{4\pi}\right)\phi_\nu\left(1 - \frac{n_j B_{ji}\psi_\nu}{n_i B_{ij}\phi_\nu}\right) = \left(\frac{n_i B_{ij} h\nu_{ij}}{4\pi}\right)\phi_\nu\left(1 - \frac{n_j g_i \psi_\nu}{n_i g_j \phi_\nu}\right) \quad (4\text{-}11)$$

and a *line source function*

$$S_l = \frac{n_j A_{ji}\psi_\nu}{n_i B_{ij}\phi_\nu - n_j B_{ji}\psi_\nu} = \left(\frac{2h\nu_{ij}^3}{c^2}\right)\left[\frac{n_i g_j \phi_\nu}{n_j g_i \psi_\nu} - 1\right]^{-1} \quad (4\text{-}12)$$

The transfer equation then reduces to the standard form [equation (2-36)].

In many cases of astrophysical interest, the simplifying approximation of *complete redistribution* is valid; then $\psi_\nu \equiv \phi_\nu$ and equations (4-11) and (4-12) reduce to

$$\chi_l(\nu) = \left(\frac{n_i B_{ij} h\nu_{ij}}{4\pi}\right)\phi_\nu\left(1 - \frac{n_j g_i}{n_i g_j}\right) \quad (4\text{-}13)$$

and

$$S_l = (2h\nu_{ij}^3/c^2)[(n_i g_j/n_j g_i) - 1]^{-1} \quad (4\text{-}14)$$

These expressions will be used through most of this book, except in Chapter 13 where a distinction between $\psi_\nu$ and $\phi_\nu$ will be made. In the case of LTE, Boltzmann statistics apply; hence $(n_j/n_i) = (n_j/n_i)^* = (g_j/g_i)\exp(-h\nu_{ij}/kT)$, and the absorption coefficient becomes

$$\chi_l^*(\nu) = (n_i B_{ij} h\nu_{ij}/4\pi)\phi_\nu[1 - \exp(-h\nu_{ij}/kT)] \quad (4\text{-}15)$$

The factor $[1 - \exp(-h\nu/kT)]$ is usually referred to as the *correction for stimulated emission*; but as is evident from equations (4-11) and (4-13), *this expression for the correction factor is valid in LTE only*. Similarly, in LTE the source function becomes

$$S_l^* = (2h\nu_{ij}^3/c^2)[\exp(h\nu_{ij}/kT) - 1]^{-1} \equiv B_\nu \quad (4\text{-}16)$$

as expected from the Kirchhoff–Planck relation [equation (2-5)].

Equation (4-14) contains implicitly the solution of the statistical equilibrium equations (cf. Chapter 5), inasmuch as it makes reference to the ratio of the populations of the levels involved. Only in the case of LTE can this ratio be expressed in terms of a single thermodynamic variable, $T$; in general it will depend upon the temperature, density, *and* the radiation field (in *all* transitions of the atom). We shall therefore refer to equation (4-14) as an *implicit form* of the source function. An alternative approach is explicitly to introduce *analytically* the solution of the statistical equilibrium equations



into the source function, yielding what we shall refer to as an *explicit form*. As we shall see in Chapters 7 and 11 through 14, the latter form is by far the more powerful and useful.

## 4-2 The Calculation of Transition Probabilities

Let us now turn to the calculation of the Einstein coefficients. Specifically, we shall derive the direct absorption probability $B_{ij}$, as $B_{ji}$ and $A_{ji}$ can be obtained from $B_{ij}$ by use of equations (4-8) and (4-9). The computation may be made on three successively more accurate levels of approximation, as follows.

(1) Classical Atom and Electromagnetic Field. The electron is considered to be a damped harmonic oscillator driven by the electromagnetic field. A *unique* absorption coefficient is derived, which is dimensionally correct and accurate for very strong lines; for weak lines it may be wrong by orders of magnitude.

(2) Quantum Mechanical Atom and Classical Electromagnetic Field. Here the correct values of $B_{ij}$ and $B_{ji}$ can be derived, but $A_{ji}$ does not appear in the formulation (although it is still given correctly by the Einstein relations).

(3) Quantum Mechanical Atom and Quantized Electromagnetic Field. Here the correct results are automatically obtained for all three coefficients, and this approach represents the complete theory.

In this section we shall carry through the calculation by method (1) for historical interest and general background, and by method (2) to obtain the correct expression for $B_{ij}$. Application of method (3) is more complicated and does not need to be carried through if one is satisfied to use the Einstein relations. A complete discussion of the third method may be found, e.g., in (**197**, Chap. 10; **293**, §§7, 14, 17; and **418**, Chap. 22).

### THE CLASSICAL OSCILLATOR

Consider first the electromagnetic radiation from a moving charge. Suppose a particle has charge $e$, velocity $\mathbf{v}$, and acceleration $\dot{\mathbf{v}}$. Then from classical electromagnetic theory (**331**, Chap. 9; **343**, Chap. 17; **494**, Chap. 20) the electric and magnetic fields at a distant position $\mathbf{r}$ relative to the charge are found to be

$$\mathbf{E}(\mathbf{r}, t) = (e\dot{v}/c^2 r)\sin\theta\,\hat{\boldsymbol{\theta}} \quad (4\text{-}17)$$

and

$$\mathbf{H}(\mathbf{r}, t) = (e\dot{v}/c^2 r)\sin\theta\,\hat{\boldsymbol{\phi}} \quad (4\text{-}18)$$

where $\theta$ denotes the angle between the acceleration $\dot{\mathbf{v}}$ and $\hat{\mathbf{r}}$, $\hat{\boldsymbol{\theta}}$ and $\hat{\boldsymbol{\phi}}$ are unit vectors in a spherical system defined by $\dot{\mathbf{v}}$ and $\hat{\mathbf{r}}$, and $\dot{v}$ is evaluated at a

time $t' = t - (r/c)$. The power radiated per cm$^2$ is given by the Poynting vector [cf. equation (1-35)]

$$\mathbf{S} = (c/4\pi)(\mathbf{E} \times \mathbf{H}) = (e^2\dot{v}^2/4\pi c^3 r^2) \sin^2\theta\, \hat{\mathbf{r}} \tag{4-19}$$

Now integrating over a sphere of radius $r$ by forming $\mathbf{S} \cdot \mathbf{dA}$ where $\mathbf{dA} = (r^2\, d\omega)\hat{\mathbf{r}} = (r^2\, d\mu\, d\phi)\hat{\mathbf{r}}$ and writing $\sin^2\theta = (1 - \mu^2)$, we find the total power radiated in all directions is

$$P(t) = (e^2\dot{v}^2/4\pi c^3) \int_0^{2\pi} d\phi \int_{-1}^{1} (1 - \mu^2)\, d\mu = (2e^2\dot{v}^2/3c^3) \tag{4-20}$$

In particular, for a *harmonic oscillator* we can write $x(t) = x_0 \cos \omega t$, $v(t) = -\omega x_0 \sin \omega t$, and $\dot{v}(t) = -\omega^2 x_0 \cos \omega t$. Substituting into equation (4-20) and averaging over one period of the oscillation by noting that $\langle \cos^2 \omega t \rangle = \frac{1}{2}$, we find

$$\langle P(\omega) \rangle = (e^2 {x_0}^2 \omega^4/3c^3) \tag{4-21}$$

Because the oscillator is radiating away energy, the oscillation will eventually decay. We may describe the decay in terms of a damping force that may be viewed as the force exerted on the moving particle by its own electromagnetic field. To calculate the effective damping force, we assume that the rate of work done by it accounts for the energy loss by the oscillator. Thus from equation (4-20) we write

$$\mathbf{F}_{\text{rad}} \cdot \mathbf{v} + (2e^2\dot{v}^2/3c^3) = 0 \tag{4-22}$$

Then

$$\int_{t_1}^{t_2} (\mathbf{F}_{\text{rad}} \cdot \mathbf{v})\, dt + (2e^2/3c^3)\left( \Big|_{t_1}^{t_2} \mathbf{v} \cdot \dot{\mathbf{v}} - \int_{t_1}^{t_2} \ddot{\mathbf{v}} \cdot \mathbf{v}\, dt \right) = 0 \tag{4-23}$$

Over a cycle the integrated term vanishes, therefore on the average

$$\mathbf{F}_{\text{rad}} = (2e^2/3c^3)\ddot{\mathbf{v}} \tag{4-24}$$

To a good order of approximation we may calculate $\ddot{\mathbf{v}}$ from its value for the undamped oscillator, namely $\ddot{\mathbf{v}} = -{\omega_0}^2\mathbf{v}$, and thus we can write

$$\mathbf{F}_{\text{rad}} = -m\gamma\mathbf{v} \tag{4-25}$$

where

$$\gamma \equiv (2e^2{\omega_0}^2/3mc^3) \tag{4-26}$$

The constant $\gamma$ is called the *classical damping constant* because of the formal resemblance of the radiation reaction term as expressed in equation (4-25) to a viscous damping term.



We now can calculate the scattering coefficient for a classical oscillator in an imposed electromagnetic field. In the classical picture the interaction is a conservative scattering process; hence we can compute the energy scattered out of a beam by calculating the energy radiated by an oscillator driven by the electromagnetic field of the incident radiation. The equation of motion for an oscillator of mass $m$ and charge $e$ driven by a field of amplitude $\mathbf{E}_0$ and frequency $\omega$ is

$$m(\ddot{\mathbf{x}} + {\omega_0}^2\mathbf{x}) = e\mathbf{E}_0 e^{i\omega t} - m\gamma\dot{\mathbf{x}} \tag{4-27}$$

Taking a trial solution for $\mathbf{x}$ that is proportional to $\exp(i\omega t)$, we find for the steady-state solution

$$\mathbf{x} = Re\left[\frac{(e/m)\mathbf{E}_0 e^{i\omega t}}{({\omega_0}^2 - \omega^2) + i\gamma\omega}\right] \tag{4-28}$$

from which we derive

$$\ddot{\mathbf{x}} = \dot{\mathbf{v}} = Re\left[\frac{-(e\omega^2/m)\mathbf{E}_0 e^{i\omega t}}{({\omega_0}^2 - \omega^2) + i\gamma\omega}\right] \tag{4-29}$$

Thus, substituting into equation (4-20) and averaging over a period, we have

$$\langle P(\omega) \rangle = \left(\frac{e^4\omega^4}{3m^2c^3}\right) \frac{{E_0}^2}{({\omega_0}^2 - \omega^2)^2 + \gamma^2\omega^2} \tag{4-30}$$

which is to be identified with the total energy scattered out of the beam. To calculate the energy scattered, we suppose that the scattering cross-section is isotropic, and write $I(\mu, \phi) = I_0\, \delta(\mu - \mu_0)\, \delta(\phi - \phi_0)$; in §1-2 we found that to produce a correspondence between the macroscopic and electromagnetic descriptions of radiation, we have $I_0 = (c{E_0}^2/8\pi)$. Thus

$$\begin{aligned}\langle P(\omega) \rangle &= \sigma(\omega) \oint I\, d\Omega = \sigma(\omega) \left(\frac{c{E_0}^2}{8\pi}\right) \int_0^{2\pi} d\phi \int_{-1}^{1} d\mu\, \delta(\mu - \mu_0)\, \delta(\phi - \phi_0) \\ &= (c{E_0}^2/8\pi)\sigma(\omega)\end{aligned} \tag{4-31}$$

Comparing equations (4-30) and (4-31) we see that the scattering coefficient is

$$\sigma(\omega) = (8\pi e^4\omega^4/3m^2c^4)[({\omega_0}^2 - \omega^2)^2 + \gamma^2\omega^2]^{-1} \tag{4-32}$$

Equation (4-32) can be simplified by noting that, because $\gamma \ll \omega$ for frequencies corresponding to light, $\sigma(\omega)$ is a sharply peaked function in the neighborhood $\omega \approx \omega_0$. To a good approximation we can replace

$$({\omega_0}^2 - \omega^2) = (\omega_0 + \omega)(\omega_0 - \omega) \approx 2\omega_0(\omega_0 - \omega)$$

and substitute equation (4-26) for $\gamma$ into equation (4-32) to find

$$\sigma(\omega) = \left(\frac{\pi e^2}{mc}\right)\left[\frac{\gamma}{(\omega - \omega_0)^2 + (\gamma/2)^2}\right] \tag{4-33}$$

The *total cross-section* can be found by integrating equation (4-33) over all frequencies, namely

$$\sigma_{\text{tot}} = \frac{\pi e^2}{mc}\int_0^\infty \frac{(\gamma/4\pi^2)\,d\nu}{(\nu - \nu_0)^2 + (\gamma/4\pi)^2} = \frac{\pi e^2}{mc}\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{dx}{1 + x^2} = \frac{\pi e^2}{mc} \tag{4-34}$$

where we have written $x = 4\pi(\nu - \nu_0)/\gamma$ and observed that, for all practical purposes, $-4\pi\nu_0/\gamma = -\infty$. The total cross-section gives a measure of the efficiency with which energy is removed from the beam. The factor in square brackets in equation (4-33) is thus a normalized profile function that is known as the *Lorentz profile* (or *damping profile*). For our present purposes, attention will be confined to the total cross-section alone; the profile function will be discussed in detail in Chapter 9.

The classical result derived above predicts a unique scattering efficiency for all transitions; this is not surprising insofar as the theory makes no reference to the actual atomic structure or to the nature of the levels between which the transition occurs. The quantum mechanical treatment shows that cross-sections for different transitions may differ by orders of magnitude. A customary way of writing the quantum mechanical result for the total cross-section is

$$\sigma_{\text{tot}} = (\pi e^2/mc) f_{ij} \tag{4-35}$$

where $f_{ij}$ is called the *oscillator strength* of the transition. In pictorial terms, $f_{ij}$ may be thought of as giving the "effective number" of classical oscillators involved in the transition; only for the strongest lines does $f_{ij}$ approach unity. The oscillator strength is related to the Einstein coefficient $B_{ij}$ by the expression

$$\sigma_{\text{tot}} = (\pi e^2/mc) f_{ij} = B_{ij}(h\nu_{ij}/4\pi) \tag{4-36}$$

### QUANTUM MECHANICAL CALCULATION

Let us now consider the calculation of $B_{ij}$ when the atom is treated according to quantum mechanics and the radiation field according to classical electrodynamics. The atomic structure is described by a *wave function* $\psi(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N, t)$ where $\mathbf{r}_1$, etc., are the positions of the bound electrons. The quantity $\psi\psi^*\,\mathbf{dr}_1 \cdots \mathbf{dr}_N$ is interpreted to be the probability of finding the atom with the electrons in the volume element $(\mathbf{r}_1, \mathbf{r}_1 + \mathbf{dr}_1)$, etc. The



wave functions are solutions of *Schrödinger's equation*

$$H\psi = i\hbar(\partial\psi/\partial t) \tag{4-37}$$

where $H$ is the total *Hamiltonian* of the system in operator form [see, e.g., (**418**, Chaps. 8–10) for a discussion of mathematical expressions for the Hamiltonian]. The Hamiltonian operator is constructed from the classical Hamiltonian according to the rule

$$H(q_i, p_i) \rightarrow H[q_i, (\hbar/i)(\partial/\partial q_i)] \tag{4-38}$$

where $q_i$ and $p_i$ are space coordinates and momenta, respectively. The atom has certain *stationary states* (or *eigenstates*) in which its energy is constant. For simplicity we shall assume that these states are *nondegenerate*. Thus if $H_A$ is the Hamiltonian of an atom which is in a stationary state $j$ of energy $E_j$, then

$$H_A\psi_j = (i\hbar)(\partial\psi_j/\partial t) \equiv E_j\psi_j \tag{4-39}$$

which implies

$$\psi_j(t) = \psi_j(0)\exp(-iE_jt/\hbar) \tag{4-40}$$

We may therefore write the general solution in the form

$$\psi_j(\mathbf{r}, t) = \phi_j(\mathbf{r})\exp(-iE_jt/\hbar) \tag{4-41}$$

where the *time-independent* solutions $\phi_j$ satisfy the equation

$$H_A\phi_j = E_j\phi_j \tag{4-42}$$

and are *orthogonal* so that

$$\int \phi_i^*\phi_j\,d\tau \equiv \langle\phi_i^*|\phi_j\rangle = \delta_{ij} \tag{4-43}$$

A general state of the system at time $t = 0$ can be expanded in terms of the eigenstates (which form a complete set) by writing

$$\psi(0) = \sum_j a_j\phi_j \tag{4-44}$$

where $a_j = \langle\phi_j^*|\psi(0)\rangle$. For a system in a general state with wavefunction $\psi(0)$, the probability of *finding* the system in a specific state $j$ by a measurement process is $a_j^*a_j = |a_j|^2$. At an arbitrary time $t$, we can write the general state as

$$\psi(t) = \sum_j a_j(t)\psi_j(t) = \sum_j a_j(t)\phi_j\exp(-iE_jt/\hbar) \tag{4-45}$$

and again the probability of finding the system in a specific state $j$ is $|a_j(t)|^2$.

If the atom is *unperturbed* (i.e., $H \equiv H_A$), then the $a_j$'s are *constant*. If, however, the atom is *perturbed* by some potential $V$, then the $a$'s will change with time; this is interpreted as the atom undergoing *transitions* from one state to another. An example of such a perturbation is that exerted by an external electromagnetic field upon the atomic electrons. In the lowest order of approximation we can assume the atom is in a uniform, time-varying electromagnetic field, $\mathbf{E} = (E_0 \cos \omega t)\mathbf{i}$. The assumption of uniformity is reasonable for light waves, which have wavelengths ($\lambda \sim 10^{-5}$ cm) that are large compared to a typical atomic dimension characterized by the *Bohr radius* ($a_0 = 5 \times 10^{-9}$ cm). The potential of the atomic electrons in the field is

$$V = e \sum_{i=1}^{N} \mathbf{E} \cdot \mathbf{r}_i \equiv \mathbf{E} \cdot \mathbf{d} = (E_0 \cos \omega t)(\mathbf{i} \cdot \mathbf{d}) \tag{4-46}$$

where $\mathbf{d}$ is the *dipole moment* of the atom. With a perturbing potential, Schrödinger's equation becomes

$$(H_A + V)\psi = i\hbar(\partial\psi/\partial t) \tag{4-47}$$

Substituting equation (4-45) for $\psi$ we have

$$(H_A + V)\sum_n a_n(t)\psi_n(t) = i\hbar \sum_n \dot{a}_n\psi_n + i\hbar \sum_n a_n(\partial\psi_n/\partial t) \tag{4-48}$$

In view of equation (4-39), equation (4-48) reduces to

$$i\hbar \sum_n \dot{a}_n\psi_n = \sum_n a_n V\psi_n \tag{4-49}$$

We may isolate a particular coefficient $\dot{a}_m$ by using the orthogonality of the $\phi$'s. Thus, multiply equation (4-49) by $\psi_m^*$ and integrate over all space. We then have

$$i\hbar \sum_n \dot{a}_n \exp[i(E_m - E_n)t/\hbar]\langle\phi_m^*|\phi_n\rangle = \sum_n a_n(t) \exp[i(E_m - E_n)t/\hbar]\langle\phi_m^*|V|\phi_n\rangle \tag{4-50}$$

Now writing $\omega_{mn} \equiv (E_m - E_n)/\hbar$ and $V_{mn} \equiv \langle\phi_m^*|V|\phi_n\rangle$, and using equation (4-43), equation (4-50) reduces to

$$\dot{a}_m(t) = (i\hbar)^{-1} \sum_n a_n(t)V_{mn}e^{i\omega_{mn}t} \tag{4-51}$$

For the perturbing potential given by equation (4-46), we see that

$$\begin{aligned} V_{mn} &= (E_0 \cos \omega t)\hat{\mathbf{i}} \cdot \langle\phi_m^*|\mathbf{d}|\phi_n\rangle = (E_0 \cos \omega t)(\hat{\mathbf{i}} \cdot \mathbf{d}_{mn}) \\ &\equiv 2h_{mn} \cos \omega t = h_{mn}(e^{i\omega t} + e^{-i\omega t}) \end{aligned} \tag{4-52}$$



The quantities $\mathbf{d}_{mn}$ are called the dipole moment *matrix elements*. Substituting equation (4-52) into equation (4-51) we have

$$\dot{a}_m(t) = (i\hbar)^{-1}\sum_n a_n(t)h_{mn}e^{i\omega_{mn}t}(e^{i\omega t} + e^{-i\omega t}) \tag{4-53}$$

We now make the simplifying assumption that, at time $t = 0$, the atom is in a definite eigenstate $k$, and we consider a time interval $T$ so short that this state is not appreciably depopulated. That is, at $t = 0$, we assume $a_k(0) = 1$ and $a_n(0) = 0$ for all $n \neq k$. Moreover, we choose $T$ such that $a_k(t) \approx 1$ for all $t \leqslant T$. Then the sum in equation (4-53) may be replaced by a single term

$$\dot{a}_m(t) = (i\hbar)^{-1}h_{mk}e^{i\omega_{mk}t}(e^{i\omega t} + e^{-i\omega t}) \tag{4-54}$$

Integrating equation (4-54) with respect to time we obtain

$$a_m(t) = \frac{h_{mk}}{i\hbar}\left\{\frac{\exp[i(\omega_{mk} - \omega)t] - 1}{(\omega_{mk} - \omega)} + \frac{\exp[i(\omega_{mk} + \omega)t] - 1}{(\omega_{mk} + \omega)}\right\} \tag{4-55}$$

As we are interested in absorption processes we choose $E_m > E_k$, so that $\omega_{mk} > 0$. From the denominator of the first term in the braces, we see that the dominant contribution to $a_m(t)$ will come when $\omega \approx \omega_{mk}$ (i.e., radiation near the line frequency is most effective in producing transitions). It is clear that the second term can be neglected in comparison with the first. Then, writing $x = (\omega - \omega_{mk})$, and forming $|a_m|^2 = a_m^* a_m$, we have

$$|a_m(t)|^2 = 4\hbar^{-2}h_{mk}{}^2x^{-2}\sin^2\left(\frac{1}{2}xt\right) = \hbar^{-2}E_0{}^2|\mathbf{i} \cdot \mathbf{d}_{mk}|^2x^{-2}\sin^2\left(\frac{1}{2}xt\right) \tag{4-56}$$

Equation (4-56) gives the number of $k \to m$ transitions (per atom in the initial state $k$) produced in time $t$ by radiation of frequency $\nu = \omega/2\pi$. To calculate the total number of transitions, we must sum over all frequencies that can contribute. Suppose that the line has a profile $\phi_\nu$ that falls sharply to zero over some characteristic frequency interval $\Delta\nu$, and that over this range (at least) the intensity of radiation (and hence $E_0{}^2$) is constant with a value $\bar{J}_\nu$. Then integrate over $d\nu = d\omega/2\pi = dx/2\pi$, and define $u \equiv \frac{1}{2}xt$, to obtain

$$\mathcal{N}_{km} = (E_0{}^2/4\pi\hbar^2)|\mathbf{i} \cdot \mathbf{d}_{mk}|^2t\int_{-U}^{U} u^{-2}\sin^2 u\, du \tag{4-57}$$

Now for thermal radiation, a characteristic frequency interval $\Delta\omega$ over which the intensity will be constant is of the order $(kT/h) \sim 10^{15}$, while transition times $t \sim 10^{-8}$ sec; hence the limits $\pm U$ on the integral may be extended formally to $\pm\infty$. The value of the integral is then found in standard tables to be $\pi$. Further, as was shown in §1-2, $E_R = (4\pi\bar{J}_\nu/c) = (E_0{}^2/8\pi)$, therefore

$$\mathcal{N}_{km} = (8\pi^2/\hbar^2c)|\mathbf{i} \cdot \mathbf{d}_{mk}|^2\bar{J}_\nu t \tag{4-58}$$

Now in terms of the Einstein coefficient $B_{km}$,

$$\mathcal{N}_{km} = B_{km}\bar{J}_\nu t \tag{4-59}$$

hence

$$B_{km} = (8\pi^2/\hbar^2 c)|\mathbf{i}\cdot\mathbf{d}_{mk}|^2 \tag{4-60}$$

In general we shall be interested in the absorptivity of bulk material. If we assume the atoms are oriented at random with respect to a beam of radiation, then $\langle|\mathbf{i}\cdot\mathbf{d}_{mk}|^2\rangle = {d_{mk}}^2\langle\cos^2\theta\rangle = \frac{1}{3}{d_{mk}}^2$, so we may write, finally

$$B_{km} = (8\pi^2\,{d_{mk}}^2/3\hbar^2 c) \tag{4-61}$$

The spontaneous emission rate follows from equations (4-8) and (4-9), which give

$$A_{mk} = (64\pi^4\nu^3/3hc^3)S(i,j) \tag{4-62}$$

In many cases the upper and lower states of a line will be degenerate (or we may wish to group several levels belonging to a *multiplet*). It is then customary to sum over all substates $k$ of the lower state $i$ and substates $m$ of the upper state $j$, and to define a *line strength* $S$ such that

$$S(i,j) \equiv \sum_{mk} {d_{mk}}^2 \tag{4-63}$$

Then we may write

$$g_j A_{ji} = (64\pi^4\nu^3/3hc^3)S(i,j) \tag{4-64}$$

or, equivalently,

$$g_i B_{ij} = (32\pi^4/3h^2 c)S(i,j) \tag{4-65}$$

and from equation (4-63)

$$g_i f_{ij} = (8\pi^2 m\nu/3he^2)S(i,j) \tag{4-66}$$

Finally, noting that $S(i,j)$ is a sum over all upper and lower substates, equation (4-66) can be used to express the total oscillator strength of a "line" connecting two degenerate levels (or of the entire multiplet connecting two sets of closely-spaced levels). Let $n'$ be the principal quantum number of the lower level, and label each sublevel with $l'$; let $n$ and $l$ correspond to the upper level. Then

$$g_{n'} f(n', n) = \sum_{l',l} g_{n'l'} f(n', l'; n, l) \tag{4-67}$$

### APPLICATION TO HYDROGEN

Hydrogen, the most abundant element in the Universe, has the simplest atomic structure, and it is possible to obtain exact analytical expressions for its wave functions and oscillator strengths. There are four quantum numbers that specify a distinct state of hydrogen: $n$, the principal quantum number, which characterizes energy; $l$, the azimuthal quantum number, which characterizes the orbital angular momentum; $m$, the magnetic quantum number, which gives the projection of the orbital angular momentum along a preferred axis (taken to be the $z$-axis); and $s$, the spin quantum number of the electron, equal to $\pm\frac{1}{2}$.



In most atomic systems, the energies of different $(n, l)$ states are distinct, but for hydrogen they depend upon the principal quantum number $n$ only, and

$$E_n = -\mathcal{R}/n^2 \tag{4-68}$$

where $\mathcal{R}$ is the *Rydberg constant*

$$\mathcal{R} = (2\pi^2\mu_H e^4/h^2) \tag{4-69}$$

Here $\mu_H$ is the *reduced mass*, given in terms of the masses of a proton, $m_p$, and electron, $m_e$, by

$${\mu_H}^{-1} \equiv {m_p}^{-1} + {m_e}^{-1} \tag{4-70}$$

The wave function has the form [cf. (**392**, Chap. 5; **418**, Chaps. 9 and 10)]

$$\psi_{nlm}(r,\theta,\phi) = R_{nl}(r)Y_l^m(\theta,\phi) \tag{4-71}$$

where $Y_l^m$ is the *spherical harmonic function* expressible in terms of associated Legendre functions, and $R_{nl}$ is the radial function, which can be expressed in terms of associated Laguerre polynomials and exponentials. These functions are normalized such that

$$\int_0^\infty {R_{nl}}^2(r)r^2\,dr = 1 \tag{4-72a}$$

and

$$\int_0^{2\pi} d\phi \int_0^\pi d\theta[Y_l^m(0,\phi)]^* {Y_{l'}}^{m'}(0,\phi)\sin 0 = \delta_{ll'}\,\delta_{mm'} \tag{4-72b}$$

In equation (4-72a), $r$ is measured in units of the *Bohr radius*

$$a_0 = h^2/(4\pi^2 e^2\mu_H) \tag{4-73}$$

It is often convenient to work with the function $P_{nl}(r) \equiv rR_{nl}(r)$ which is defined such that ${P_{nl}}^2$ measures the charge density of the electronic wave function.

Because all states with a given $n$ are degenerate, we require the statistical weight $g_n$ and normally will work with the oscillator strength $f(n', n)$ for *all* transitions $(n' \to n)$. The statistical weight is

$$g_n = 2n^2 \tag{4-74}$$

which follows from the fact that the allowed values of $l$ are $0 \leqslant l \leqslant n-1$; those of $m$ are $-l \leqslant m \leqslant l$; and each $nlm$ state has two possible spin orientations $s = \pm\frac{1}{2}$.

*Exercise 4-2*: Derive equation (4-74).

Because the wave functions are known analytically, explicit expressions can be derived for the oscillator strengths:

$$f(n', l'; n, l) = \frac{1}{3}\left(\frac{1}{n'^2} - \frac{1}{n^2}\right)\frac{\max(l, l')}{(2l'+1)}\sigma^2(n', l'; n, l) \tag{4-75}$$

and

$$f(n', n) = \frac{1}{3n'^2}\left(\frac{1}{n'^2} - \frac{1}{n^2}\right)\left[\sum_{l'=1}^{n'-1} l'\sigma^2(n', l'; n, l'-1) + \sum_{l'=0}^{n'-1}(l'+1)\sigma^2(n', l'; n, l'+1)\right] \tag{4-76}$$

where

$$\sigma^2(n', l'; n, l) \equiv \left(\int_0^\infty P_{n'l'}(r)P_{nl}(r)r\,dr\right)^2 \tag{4-77}$$

An explicit expression for $\sigma^2$ was first derived by Gordon (**254**) and an explicit form for $f(n', n)$ was derived by Menzel and Pekeris (**417**). Extensive tables of $f(n', n)$ can be found in (**417**) and (**257**). A very convenient form for hydrogen oscillator strengths is obtained by expressing them in terms of the semiclassical value derived by Kramers (**363**), namely

$$f_K(n', n) = \frac{32}{3\pi\sqrt{3}}\left(\frac{1}{n'^2} - \frac{1}{n^2}\right)^{-3}\left(\frac{1}{n^3n'^5}\right) \tag{4-78}$$

which shows the principal dependences of $f$ upon $n'$ and $n$. It is then customary to express the *exact* $f$-value in terms of Kramers' approximation $f_K$ by writing

$$f(n', n) = g_I(n', n)f_K(n', n) \tag{4-79}$$

where $g_I(n', n)$ is called the *Gaunt factor*. The Gaunt factors are all numbers of order unity; an extensive tabulation of $g_I(n', n)$ can be found in (**60**).

*Exercise 4-3*: Using the analytical expressions for hydrogen wave functions given in texts on quantum mechanics [e.g., (**392**, 183)], calculate the $f$-values for $L\alpha(n' = 1 \rightarrow n = 2)$ and $H\alpha(n' = 2 \rightarrow n = 3)$. Obtain values for each $f(n', l'; n, l)$ and combine these to find $f(n', n)$. Compare your values with those given in tables [e.g., (**9**, 70)].



TRANSITION PROBABILITIES FOR LIGHT ELEMENTS

(a) *Hartree–Fock Method.* When the atom has more than one electron, the wave equation can no longer be solved in closed form, and approximations must be made. The actual Hamiltonian for an $N$-electron atom is

$$H = -(\hbar^2/2m)\sum_{i=1}^{N}\nabla_i^2 - \sum_{i=1}^{N}(Ze^2/r_i) + \sum_{\substack{\text{all pairs}\\(i,j)}}(e^2/|\mathbf{r}_i - \mathbf{r}_j|) \tag{4-80}$$

The first term represents the kinetic energy of the electrons, the second their electrostatic potential with the nucleus of charge $Z$, and the third their mutual Coulomb repulsion. It is the last term that causes the principal difficulties.

One of the most important methods of deriving approximate wave functions is *Hartree's self-consistent field method.* In this approach, the sum over electron pairs is replaced for each electron by its spherical average. An excellent description of how this average is computed is given in (**576**, Chaps. 3 and 9). Each electron then moves in a potential that depends only upon its distance from the nucleus, and we make the replacement

$$\sum_{\substack{\text{all pairs}\\(i,j)}}(e^2/|\mathbf{r}_i - \mathbf{r}_j|) \rightarrow \sum_i V_i(r_i) \tag{4-81}$$

This results in the approximation of the actual potential by a *central field.* With a central field potential, the angular factors in the Schrödinger equation can be separated out in exactly the same way as for hydrogen, and for each electron the wave function has the form

$$U_i(r, \theta, \phi; n, l, m, s) = r^{-1}P_{nl}(r)Y_l^m(\theta, \phi)X(s) \tag{4-82}$$

where the normalizations given in equation (4-72) still apply. The functions $U_i$ are called *electron orbitals.* The radial equation for each orbital is of the form ($r$ in units of $a_0$, $E$ in Rydbergs)

$$(d^2P_{nl}/dr^2) + [E_{nl} + 2r^{-1}Z_{\text{eff}}(r) - l(l+1)r^{-2}]P_{nl} = 0 \tag{4-83}$$

Here $Z_{\text{eff}}(r)$ is the "effective nuclear charge" sensed by an electron after allowance is made for shielding by other electrons [using the central fields of equation (4-81)]. The atom is now considered to be made up of $N$ such orbitals, and these are used to construct the wave function for the entire configuration.

Because of the *Pauli exclusion principle*, the set of four quantum numbers $(n, l, m, s)$ for each orbital cannot be identical for any two orbitals. Also, the

wave function of the atom must be constructed so that it is *antisymmetric* under the interchange of the coordinates of any two electrons. In practice these conditions may be met by writing the wave function as a *Slater determinant* (**576**, Chap. 12),

$$\phi(\mathbf{r}_1, \ldots, \mathbf{r}_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} U_1(\alpha) & U_1(\beta) & \cdots & U_1(\nu) \\ U_2(\alpha) & U_2(\beta) & \cdots & U_2(\nu) \\ \vdots & \vdots & & \vdots \\ U_N(\alpha) & U_N(\beta) & \cdots & U_N(\nu) \end{vmatrix} \tag{4-84}$$

where the numbers $1, 2, \ldots, N$ denote the orbitals of electrons 1, 2, etc., while $\alpha, \beta, \ldots, \nu$ stand for the space and spin coordinates of electrons $\alpha, \ldots, \nu$, respectively.

The solution for the wave function is carried out iteratively. Thus $Z_{\text{eff}}(r)$ depends in an involved way on integrals over the electron orbitals, but in turn it determines those orbitals. Therefore, we start with an approximate set of orbitals, compute $Z_{\text{eff}}$, solve for the $P_{nl}$'s, recompute $Z_{\text{eff}}$, and iterate until the procedure converges. The calculations are time-consuming and laborious, but are within the capabilities of modern computers, and a large number of wave functions for a wide variety of atomic configurations are now available.

A specific term in an atomic spectrum can be characterized by certain quantum numbers describing the atom as a whole. In light atoms these describe the total orbital angular momentum **L** (the vector sum of the individual $\mathbf{l}_i$'s), the total spin angular momentum **S** (the vector sum of the $\mathbf{s}_i$'s), and the total angular momentum **J**, which is the vector sum of **L** and **S**. This type of coupling of the individual momenta is called ($L$–$S$) or *Russell–Saunders coupling*. As a given **L**, **S**, and **J** may result from more than one arrangement of the individual **l**'s, **m**'s, and **s**'s of the orbitals, the complete wave function will, in general, consist of a *sum* of Slater determinants, and thus may be very complicated.

In calculating transition probabilities, it is generally assumed that only one orbital is different between the initial and final state—i.e., only one electron undergoes a transition. In this case the matrix element $r_{ij}$ can be split into factors, one coming from the initial and final *radial* wave functions, and another depending on the *angular* and *spin* wave functions. It is customary, therefore, to write the expression for the line strength in the form

$$S(n', L', S', J'; n, L, S, J) = a_0{}^2 e^2 \sigma^2(n', l'; n, l) \mathscr{S}(\mathscr{M}) \mathscr{S}(\mathscr{L}) \tag{4-85}$$

Here

$$\sigma^2 \equiv (4l_{\max}^2 - 1)^{-1} \left( \int_0^\infty P_{n'l'} P_{nl} r \, dr \right)^2 \tag{4-86}$$



where $l_{\max} = \max(l, l')$. The factor $\mathscr{S}(\mathscr{M})$ is the *strength of the multiplet*, depending on $nLS$ and $n'L'S'$, and the factor $\mathscr{S}(\mathscr{L})$ is the *strength of the line within the multiplet*. Extensive tables of $\mathscr{S}(\mathscr{M})$ and $\mathscr{S}(\mathscr{L})$ can be found in (**11**, Chap. 8 and Appendix; **9**, §§26–28; **250**; **251**), and general formulae for computing these factors are given in (**534**; **535**; **572**, §§10.8–10.10; and **191**). Generally, the most difficult part of the calculation is the determination of $\sigma^2$, but serious complications also occur when there are deviations from $L$–$S$ coupling.

(b) *The Coulomb Approximation.* Because of the labor involved in obtaining $\sigma^2$ from Hartree–Fock calculations, it is desirable to have an approximate method that can be applied easily. Such an approach was developed by Bates and Damgaard (**74**), who pointed out that often the largest contribution to the radial integral comes from large values of $r$, where the electron moves in a nearly Coulomb potential. In this event, the integral can be approximated using hydrogenic wave functions, provided the principal quantum numbers are chosen to give the observed energy of the level. If $Z$ is the charge in the asymptotic potential, then the appropriate effective quantum number is $n_l^* = Z/\varepsilon_{nl}{}^{\frac{1}{2}}$ where $\varepsilon_{nl}$ is the energy of the level below the continuum, measured in Rydbergs; in general, $n_l^*$ is *not* an integer. Bates and Damgaard then show that one can write

$$\sigma(n_{l-1}^*, l - 1; n_l^*, l) = \mathscr{F}(n_l^*, l) \mathscr{I}(n_{l-1}^*, n_l^*, l)/Z \tag{4-87}$$

Extensive tables of the functions $\mathscr{F}$ and $\mathscr{I}$ can be found in (**74**) and (**483**); an extension of the theory is given in (**389**). Because of the simplicity of the method, it has been widely applied in astrophysical analyses; an extensive tabulation of Coulomb approximation $f$-values is given in (**264**, 363–441).

(c) *Experimental Methods.* In many cases the Coulomb approximation is inaccurate, while a more accurate quantum mechanical calculation is simply too complicated to carry out. In these cases the $f$-values must be determined by experiments, which in addition, provide a direct comparison standard to test the accuracy of various theoretical computations. A wide variety of experimental techniques exist; brief descriptions of some of the more useful methods are given in (**11**, 300–310; **261**, 146–149; **264**, Chap. 15).

There is a huge literature containing both experimental and theoretical determinations of $f$-values or transition probabilities; a complete bibliography of this work is given in (**454**; **230**; and **231**). Compilations of critically evaluated (i.e., "best values") transition probabilities for many of the elements of astrophysical interest are given in (**453**; **584**; **670**; **672**; and **673**).

*Exercise 4-4*: Calculate $f$-values for the He I $\lambda$5876 ($2p\ ^3P$–$3d\ ^3D$), He I $\lambda$6678 ($2p\ ^1P$–$3d\ ^1D$), and He I $\lambda$4471 ($2p\ ^3P$–$4d\ ^3D$) lines using equations (4-66) and (4-85) with Coulomb approximation values of $\sigma^2$ and multiplet- and line-strengths from tables. Compare with standard values in (**672**).

## 4-3 The Einstein–Milne Relations for the Continuum

The Einstein relations were generalized to bound–free processes by Milne (**461**) in a paper of considerable interest and importance. We consider photoionization processes that start with atoms (or ions) in a definite bound state (not necessarily the ground state) and produce an ion in a definite state (perhaps excited) of the next highest ionization stage plus a free electron moving with velocity $v$. The inverse process is a recombination of a free electron by a collision with an ion (in the particular state mentioned above) to form an atom (in the proper state). The recombination process can occur *spontaneously* or can be *induced* by incident radiation. Let $n_0$ be the number density of the atoms, $n_1$ the density of the ions, and $n_e$ the density of free electrons. The electrons have a Maxwellian velocity distribution, and we write $n_e(v)\,dv$ for the number with speeds on the range $(v, v + dv)$. Let $p_\nu$ be the *probability of photoionization* of an atom by a photon in the frequency range $(\nu, \nu + d\nu)$; then the *number of photoionizations* in time $dt$ on this frequency range is $n_0 p_\nu I_\nu\,d\nu\,dt$. The usual energy *absorption coefficient* $\alpha_\nu$ is related to $p_\nu$ by the expression $\alpha_\nu \equiv p_\nu h\nu$. Further, let $F(v)$ be the *spontaneous recapture probability* and $G(v)$ be the *induced recapture probability* for electrons in the velocity range $(v, v + dv)$ by the ions; then the *number of recombinations* by electrons of velocity $v$ in time $dt$ is

$$n_1 n_e(v)[F(v) + G(v)I_\nu]v\,dv\,dt$$

The energy of the photon required to ionize the atom (and thus of the photon emitted in the recombination process) is

$$h\nu = \chi_I + \frac{1}{2}mv^2 \tag{4-88}$$

where $\chi_I$ is the ionization potential from the atomic to the ionic state (i.e., the energy difference between these states).

Now *in thermodynamic equilibrium, the number of photoionizations must exactly equal the number of recombinations.* In T.E., $I_\nu \equiv B_\nu$, hence

$$n_0^* p_\nu B_\nu = n_1^* n_e(v)[F(v) + G(v)B_\nu](h/m) \tag{4-89}$$

where the asterisks denote T.E. occupation numbers, and equation (4-88) has been used to write $h\,d\nu = mv\,dv$. Solving for $B_\nu$ we find

$$B_\nu = [F(v)/G(v)]\{[n_0^* p_\nu m/n_1^* n_e(v)hG(v)] - 1\}^{-1} \tag{4-90}$$

This expression is to be compared with the standard expression for the Planck function, namely $B_\nu(T) = (2h\nu^3/c^2)[e^{h\nu/kT} - 1]^{-1}$. For a Maxwellian velocity distribution [cf. equation (5-2)]



$$n_e(v)\,dv = n_e(m/2\pi kT)^{\frac{3}{2}} \exp(-mv^2/2kT)4\pi v^2\,dv \tag{4-91}$$

Anticipating the results of Chapter 5 [cf. equation (5-14)], the T.E. relation between $n_0^*$ and $n_1^*$ is

$$(n_0/n_1)^* = n_e(g_0/2g_1)(h^2/2\pi mkT)^{\frac{3}{2}} \exp(\chi_I/kT) \tag{4-92}$$

Using equations (4-88), (4-91) and (4-92) we find that

$$n_0^* p_\nu m/n_1^* n_e(v)hG(v) = (h^2 g_0/8\pi m^2 g_1 v^2)[p_\nu/G(v)]e^{h\nu/kT} \tag{4-93}$$

Thus to reduce equation (4-90) to the Planck function we must have

$$F(v) = (2h\nu^3/c^2)G(v) \tag{4-94}$$

and

$$p_\nu = (8\pi m^2 v^2 g_1/h^2 g_0)G(v) = (4\pi c^2 m^2 v^2 g_1/h^3 g_0 \nu^3)F(v) \tag{4-95}$$

These are the continuum analogues of equations (4-8) and (4-9). Again we recognize that, although these relations have been derived from thermodynamic equilibrium arguments, the quantities $p_\nu$, $F(v)$, and $G(v)$ must really depend only on the properties of the atom; hence equations (4-94) and (4-95) are true in general.

The great importance of the results just derived becomes more clearly manifest when we write the transfer equation assuming that at the frequency under consideration only the particular photoionization and recombination processes considered above occur. The generalization to a multilevel, multiatom case with several overlapping opacities and emissivities is trivial because each term adds *linearly* and the conclusions we shall derive apply to the sum. The transfer equation is

$$\mu(\partial I_\nu/\partial z) = -n_0 p_\nu h\nu I_\nu + n_1 n_e(v)[F(v) + G(v)I_\nu](h^2\nu/m) \tag{4-96}$$

Neither $n_0$ nor $n_1$ necessarily has its LTE value in the above equation. If we are to write the transfer equation in standard form, then it is clear that the *absorption coefficient corrected for stimulated emission* must be

$$\kappa_\nu = \{n_0 - n_1 n_e(v)[hG(v)/mp_\nu]\}(p_\nu h\nu) \tag{4-97}$$

Using equations (4-88), (4-91), (4-92) and (4-95), and recalling that $\alpha_\nu = p_\nu h\nu$, we find

$$\kappa_\nu = (n_0 - n_0^* e^{-h\nu/kT})\alpha_\nu \tag{4-98}$$

In equation (4-98), $n_0^*$ denotes the LTE value of $n_0$ computed from equation (4-92) using the *actual* values of $n_1$ and $n_e$ (i.e., the LTE population relative

to the actual ion density). In the particular case of LTE where $n_0 \equiv n_0^*$,

$$\kappa_\nu^* = n_0^* \alpha_\nu (1 - e^{-h\nu/kT}) \tag{4-99}$$

As was true for bound–bound transitions, the term $(1 - e^{-h\nu/kT})$ is usually called the correction factor for stimulated emission; but it is clear that this expression is correct *only* for LTE. Indeed we see from equation (4-98) that the stimulated emission *always* occurs at the LTE rate (if we understand $n_0^*$ to have the meaning given above); this *must* be true because the recombination process is a *collisional* process involving particles with an *equilibrium* (i.e., Maxwellian) velocity distribution. Note the contrast here with the result given in equation (4-13) for bound–bound transitions, where in general the stimulated emission term does *not* have its equilibrium value. When departures from LTE affect the bound–free opacity, they change the *direct* absorption term involving $n_0$ (which in general will not equal $n_0^*$). We shall use these results both in calculating the stimulated emission rates in the equations of statistical equilibrium [cf. equation (5-63)] and in writing a general expression for the opacity [cf. equation (7-1)].

Returning to equation (4-96) and examining the term involving $F(v)$, it is clear that the *emissivity* is

$$\eta_\nu = [h n_1 n_e(v) F(v)/m p_\nu] \alpha_\nu \tag{4-100}$$

which, with the help of equations (4-88), (4-91), (4-92), and (4-95), can be written

$$\eta_\nu = (2h\nu^3/c^2) n_0^* \alpha_\nu e^{-h\nu/kT} = n_0^* \alpha_\nu (1 - e^{-h\nu/kT}) B_\nu = \kappa_\nu^* B_\nu(T) \tag{4-101}$$

Thus the continuum emissivity always occurs at the LTE rate (if $n_0^*$ is defined as above), which is what we would expect, for the recombination process is collisional. Notice that this derivation recovers the Kirchhoff–Planck law, equation (2-6), and extends its validity somewhat. Again, notice the contrast with the bound–bound spontaneous emission where departures from LTE enter directly if $n_j$ is not identical to $n_j^*$. These results will be exploited in calculating spontaneous emission rates in the statistical equilibrium equations [cf. equation (5-61)] and in writing a general expression for the emissivity [cf. equation (7-2)].

*Exercise 4-5*: Verify equations (4-93), (4-98), and (4-101).

## 4-4 Continuum Absorption Cross-Sections

Cross-sections for bound–free absorption can be calculated quantum mechanically by essentially the same methods as used in §4-2 for bound–



bound transitions. Consider absorptions from a bound state $n$, of statistical weight $g_n$, to the continuum in a frequency interval $\Delta\nu$. The free states have wave functions characterized by $E$, the energy of the free electron, and are normalized such that

$$\langle E'|E\rangle = \delta(E' - E) \tag{4-102}$$

so that there are $\Delta E$ states in the energy interval $\Delta E$. Thus by analogy with equation (4-65) and in view of (4-36) we can write

$$g_n \alpha_\nu \, \Delta\nu = (8\pi^2/3h^2c) \, \Delta E \langle E|\mathbf{d}|n\rangle^2 (h\nu/4\pi) \tag{4-103}$$

or

$$\alpha_\nu = (8\pi^3\nu/3cg_n)\langle E|\mathbf{d}|n\rangle^2 \tag{4-104}$$

The calculation of free-state wave functions will not be considered in this book; the reader should refer to standard texts on quantum mechanics—e.g., (**197**) or (**418**)—for information on this subject. Further, we shall not consider the details of calculations based on equation (4-104), though results of such work will be quoted. An approximate method for the evaluation of $\alpha_\nu$ by means of equation (4-104), the quantum defect method, is described below.

An alternative formula for $\alpha_\nu$ can be written if we suppose each continuum state $k$ to have an effective oscillator strength $f_{nk}$ for absorptions from the bound state $n$. If there are $\Delta k$ free states in the frequency interval $\Delta\nu$, then

$$\alpha_\nu = (\pi e^2/mc) f_{nk} (\Delta k/\Delta\nu) \tag{4-105}$$

This formulation is useful for calculating the cross-sections of hydrogen.

The *quantum defect method*, developed by Seaton and Burgess (**566**; **120**), is the continuum analogue of the Coulomb approximation. This method exploits the fact that the dominant contribution to the matrix element $\langle E|\mathbf{d}|n\rangle^2$ often occurs in regions where the wave functions can be represented by Coulomb wave functions in the appropriate potential. Consider absorptions from a bound state $(n, l)$ to the continua $(E, l \pm 1)$ where $E$ is the energy of the free electron. Let $I_{nl}$ be the ionization energy of this state, expressed in Rydbergs, and let $Z$ be the charge on the ion after the electron is removed. Then define the effective quantum number $\nu_{nl}$ such that $I_{nl} = Z^2/\nu_{nl}^2$. In general, $\nu_{nl}$ will not be equal to the principal quantum number $n$ of the shell to which the electron belongs, and we can define a *quantum defect* $\mu(\nu, l) \equiv n - \nu_{nl}$. The quantum defect can be found for each level $(nlSL)$ of a given spectroscopic type, defined by $(lSL)$ (e.g., $^3P$ or $^4D$), in a series. Defining $\varepsilon_{nl} \equiv -1/\nu_{nl}^2$, we can determine the behavior of $\mu(\varepsilon_{nl}, l)$ versus $\varepsilon_{nl}$; in favorable cases $\mu$ is a simple function of $\varepsilon$ (say, constant, or linear in $\varepsilon$). It is then assumed that this variation of $\mu$ with $\varepsilon$ can be extrapolated into the continuum (i.e., for $\varepsilon > 0$) to give $\mu'(\varepsilon)$. This establishes the properties of the

continuum wave functions. The radial matrix element can then be evaluated using hydrogenic wave functions, and the cross-section when the energy of the ejected electron is $k^2 = Z^2\varepsilon$ (in Rydbergs) can be written

$$\alpha(nl, k^2) = 8.56 \times 10^{-19}[(I_{nl} + k^2)/I_{nl}{}^2] \sum_{l'=l\pm1} C_{l'}|g(vl; \varepsilon l')|^2 \text{ cm}^2 \quad (4\text{-}106)$$

$$\text{Here } g(vl; \varepsilon l') = [\zeta(v, l)]^{-\frac{1}{3}} G(vl; \varepsilon l') \cos\{\pi[v + \mu'(\varepsilon) + \chi(vl; \varepsilon l')]\} \quad (4\text{-}107)$$

$$\zeta(v, l) \equiv 1 + 2v^{-3}[\partial\mu(\varepsilon)/\partial\varepsilon] \quad (4\text{-}108)$$

and $G(vl; \varepsilon l')$ and $\chi(vl; \varepsilon l)$ are tabulated functions (**503**) [NOTE: the notation in the reference cited differs from that in (**120**), which is used here]. The coefficients $C_{l'}$ are algebraic factors obtained from the integrations over angular and spin coordinates and are tabulated in (**120**) for several important cases; they are analogues of the factor $\mathscr{S}(\mathscr{M})\mathscr{S}(\mathscr{L})$ appearing in equation (4-85) for bound–bound oscillator strengths.

The quantum defect method, despite its simplicity, often gives very good cross-sections [see (**120**; **503**)] and has been widely employed in astrophysical work. A number of quantum defects $\mu(\varepsilon)$ are given in (**503**) where they are used to calculate cross-sections and opacities for abundant elements in stellar atmospheres [see also (**502**)]. For brevity, only absorption by hydrogen and helium, and their ions, will be discussed in this chapter; these are the most abundant elements in stellar material, and usually dominate the opacity. References to other opacity sources will be given in §7-2.

### HYDROGEN

A simple way of obtaining bound–free and free–free absorption cross-sections for hydrogen was suggested by Menzel and Pekeris (**417**). They introduced the formalism of representing *bound* states by *real* (integer) quantum numbers, and *free* states by imaginary quantum numbers. The bound-state energies relative to the continuum are given by equation (4-68), and it follows that the energy of the transition $(n' \to n)$ is

$$h\nu_{n'n} = \mathscr{R}[(1/n')^2 - (1/n)^2] \quad (4\text{-}109)$$

If a free state has the imaginary quantum number $ik$, then by analogy

$$h\nu_{n'k} = \mathscr{R}[(1/n')^2 + (1/k)^2] = (\mathscr{R}/n'^2) + \frac{1}{2}mv^2 \quad (4\text{-}110)$$

where the first term clearly represents the ionization potential from bound state $n'$, and the second the energy of the free electron. Note that $k \to \infty$ at the ionization limit and becomes small high in the continuum.



The formula for the continuum oscillator strength follows from a generalization of equations (4-78) and (4-79) to

$$f_{n'k} = \left(\frac{32}{3\pi\sqrt{3}}\right)\left(\frac{1}{n'^5k^3}\right)\left(\frac{1}{n'^2} + \frac{1}{k^2}\right)^{-3} g_{\text{II}}(n', k) \quad (4\text{-}111)$$

where $g_{\text{II}}$ is the bound–free Gaunt factor. Formulae for the Gaunt factor are given in (**417**) and an extensive numerical tabulation is given in (**352**); $g_{\text{II}}$ is a number of order unity at the ionization threshold, shows a slow rise to about 1.10 (in the limit as $n' \to \infty$) at about 1 Rydberg above threshold, and then decreases to small values in the $X$-ray region. The absorption cross-section can now be derived by substituting equation (4-111) into (4-105), noting from equation (4-110) that for $n'$ fixed,

$$(dk/dv) = -(hk^3/2\mathscr{R}) \quad (4\text{-}112)$$

We then find

$$\alpha_\nu = \left(\frac{\pi e^2}{mc}\right)\left(\frac{hk^3}{2\mathscr{R}}\right)\left(\frac{32}{3\pi\sqrt{3}}\right)\left(\frac{1}{n'^5k^3}\right)\frac{g_{\text{II}}(n', k)}{(h\nu/\mathscr{R})^3} \quad (4\text{-}113)$$

which, in view of equation (4-69), reduces to

$$\alpha_\nu = \left(\frac{64\pi^4me^{10}}{3\sqrt{3}\,ch^6}\right)\frac{1}{n'^5\nu^3}\,g_{\text{II}}(n', \nu) = \mathscr{K}\frac{g_{\text{II}}(n', \nu)}{n'^5\nu^3} \quad (4\text{-}114)$$

where $\mathscr{K} = 2.815 \times 10^{29}$. Thus bound–free absorption from level $n$ commences abruptly at the threshold frequency $\nu_n = (\mathscr{R}/hn^2)$ and falls off at higher frequencies as $\nu^{-3}$ (neglecting the weak variation of the Gaunt factor).

The threshold cross-section is given by

$$\alpha(\nu_n, n) = 7.91 \times 10^{-18}\, n g_{\text{II}}(n, \nu_n) \text{ cm}^2$$

The opacity per cm$^3$ of the stellar material can be computed by multiplying the cross-section for level $n$ by the number of hydrogen atoms (per cm$^3$) in that level, and summing over all levels that can absorb at a given frequency $\nu$ (i.e., all $n$ such that $\nu_n \leqslant \nu$). The bound–free opacity of hydrogen calculated in this way has a jagged character, as shown in Figure 4-1. Except for the hottest stars, most of the hydrogen is in the ground state, and the absorption edge at $\lambda$912 Å (one Rydberg) is extremely strong. For 912 Å $\leqslant \lambda \leqslant$ 3647 Å, absorptions from the ground state can no longer occur, and the dominant opacity source is photoionization from the $n = 2$ level (Balmer continuum). Similarly, for 3647 Å $\leqslant \lambda \leqslant$ 8206 Å, both $n = 1$ and $n = 2$ cannot absorb, and the dominant continuum is from $n = 3$ (Paschen continuum); and so on. Actually the opacity variation shown in Figure 4-1 is idealized in that there

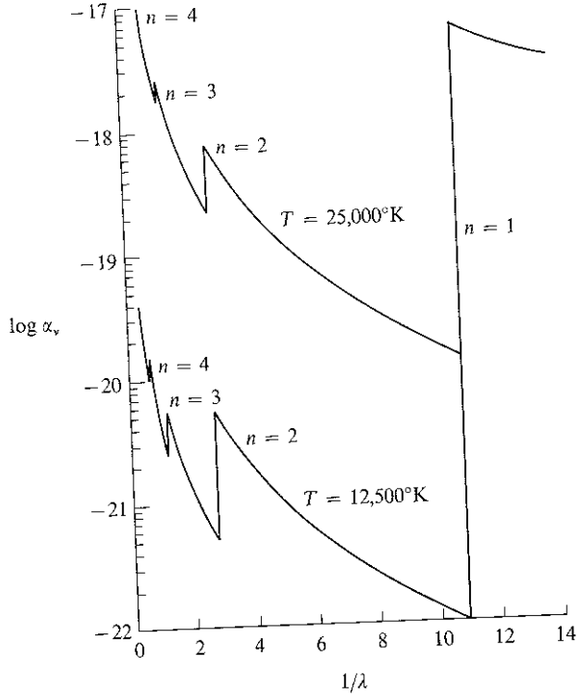



FIGURE 4-1
Opacity from neutral hydrogen at $T = 12{,}500°$K and $T = 25{,}000°$K, in LTE; photoionization edges are labeled with the quantum number of state from which they arise. *Ordinate*: sum of bound–free and free–free opacity in cm$^2$/atom; *abscissa*: $1/\lambda$ where $\lambda$ is in microns.

exists a series of lines converging on each photoionization threshold at the series limit. Near the limit the lines blend together smoothly and merge into the continuum. Bound–free absorption from hydrogen is the dominant continuum opacity source in stars of spectral types *A* and *B*.

Let us now consider the free–free opacity of hydrogen. In this process, a free electron passing near a proton causes a transitory dipole moment, and absorptions and emissions of photons (with a consequent change in the electron's energy) become possible. By analogy with the calculation of bound–free absorption, we introduce imaginary quantum numbers for *both* the initial and final states, say *ik* and *il*, such that, if $v$ is the initial velocity of the free electron, and $\nu$ is the frequency of the radiation absorbed, then

$$\mathscr{R}k^{-2} = \frac{1}{2}mv^2 \tag{4-115}$$



and

$$\mathscr{R}k^{-2} + h\nu = \mathscr{R}l^{-2} \tag{4-116}$$

Assume that absorptions take place from a band of states $dk$ into a band of states $dl = (dl/dv)\,\Delta v$; then in equation (4-105) we replace $f_{nk}$ with $f_{kl}\,dk$, and $\Delta k$ with $dl$ to obtain

$$\alpha(\nu, v) = (\pi e^2/mc) f_{kl}\,dk(dl/dv) \tag{4-117}$$

as the absorption coefficient *per ion and per electron moving with velocity v*. The appropriate generalization of equations (4-78) and (4-79) is

$$f_{kl} = \frac{64}{3\pi\sqrt{3}}\frac{1}{g_k}\left(\frac{1}{k^2} - \frac{1}{l^2}\right)^{-3}\frac{g_{\text{III}}(k, l)}{k^3 l^3} \tag{4-118}$$

where $g_k$ is the statistical weight of a free electron, given by quantum statistics as

$$g_k = (2h^{-3})(4\pi m^3 v^2\,dv) = (16\pi\mathscr{R}m^2 v/h^3k^3)\,dk \tag{4-119}$$

the second equality following from equation (4-115). Substituting into equation (4-117) we find

$$\alpha(\nu, v) = \left(\frac{\pi e^2}{mc}\right)\left(\frac{64}{3\pi\sqrt{3}}\right)\left(\frac{h^3k^3}{16\pi\mathscr{R}m^2 v}\right)\left(\frac{\mathscr{R}}{h\nu}\right)^3\frac{g_{\text{III}}(\nu, v)}{k^3 l^3}\left(\frac{dl}{dv}\right) \tag{4-120}$$

and making use of the relation that for $k$ (or $v$) fixed, $(dl/dv) = (hl^3/2\mathscr{R})$ from equation (4-116), we obtain

$$\alpha(\nu, v) = \left(\frac{2\mathscr{R}he^2}{3\pi\sqrt{3}\,m^3c}\right)\left(\frac{g_{\text{III}}(\nu, v)}{\nu^3 v}\right) \tag{4-121}$$

The total absorption cross-section per ion and per electron is obtained by summing over all incident electron velocities, assuming a Maxwellian velocity distribution as given by equation (4-91). The result is

$$\alpha(\nu, T) = \left(\frac{4e^6}{3ch}\right)\left(\frac{2\pi}{3\,km^3}\right)^{\frac{1}{2}} T^{-\frac{1}{2}}\nu^{-3}\bar{g}_{\text{III}}(\nu, T) \tag{4-122}$$

where use has been made of equation (4-69), and $\bar{g}_{\text{III}}$ is the *thermal average* of the Gaunt factor

$$\bar{g}_{\text{III}}(\nu, T) \equiv \int_0^\infty g_{\text{III}}(\nu, v)e^{-u}\,du \tag{4-123}$$

where $u \equiv (mv^2/2kT)$.

*Exercise 4-6*: Verify equations (4-122) and (4-123).

Inserting numerical values for the atomic constants, multiplying by the electron and proton densities, and correcting for stimulated emission (notice that because the process is collisional it is always in LTE, at the *actual* electron and ion densities), we obtain the opacity coefficient

$$\kappa_\nu(\text{free–free}) = 3.69 \times 10^8\, \bar{g}_{\text{III}}(\nu, T)\nu^{-3}T^{-\frac{1}{2}}n_e n_p(1 - e^{-h\nu/kT}) \quad (4\text{-}124)$$

Formulae for $\bar{g}_{\text{III}}$ are given in (**417**) and extensive tables can be found in (**85**) and (**352**). The free–free opacity plays an ever more important role at low frequencies compared to the bound–free, because of the decreasing number of photoionization edges that contribute as $\nu \to 0$. Further, the free–free becomes more important at high temperatures, for as can be seen from equation (4-92), in the limit $(kT/\chi_{\text{ion}}) \gg 1$, the bound state populations vary as $n_i \propto n_e n_p T^{-\frac{3}{2}}$; hence the ratio of free–free to bound–free opacity rises $\propto T$ in the high-temperature limit. The free–free process is the dominant true absorption mechanism in, e.g., the O-stars.

### THE NEGATIVE HYDROGEN ION

Hydrogen, because of its large polarizability, can form a *negative ion* consisting of a proton and two electrons. This ion has a single bound state with a binding energy of 0.754 eV. Because of its low binding energy, $H^-$ does not exist at high temperatures (it is destroyed by ionization) but is prevalent mainly in the atmospheres of solar-type and cooler stars. It was recognized by Pannekoek and Wildt that $H^-$ could be an important opacity source in such stars. As it turns out, the abosprtion cross-section of $H^-$ is large and, although only a small fraction of the hydrogen exists in this form, the opacity from $H^-$ is the *dominant* one in the atmospheres of cooler stars.

The negative hydrogen ion can absorb and emit radiation via both bound–free and free–free processes; i.e.,

$$H^- + h\nu \rightleftharpoons H + e(v) \quad (4\text{-}125)$$

where $\frac{1}{2}mv^2 = h\nu - 0.754$ eV, and

$$H + e(v) + h\nu \rightleftharpoons H + e(v') \quad (4\text{-}126)$$

where $\frac{1}{2}mv'^2 = \frac{1}{2}mv^2 + h\nu$. In the free–free process, an electron passing near to a neutral hydrogen atom induces, by polarization, a temporary dipole moment that can interact with the radiation field, leading to absorptions and emissions. The bound–free absorption process has its threshold at about 16500 Å (1.65 $\mu$), corresponding to the detachment energy. It reaches a maximum cross-section of about $4 \times 10^{-17}$ cm$^2$ at 8500 Å and decreases toward shorter wavelengths. The free–free cross-section is about equal to the



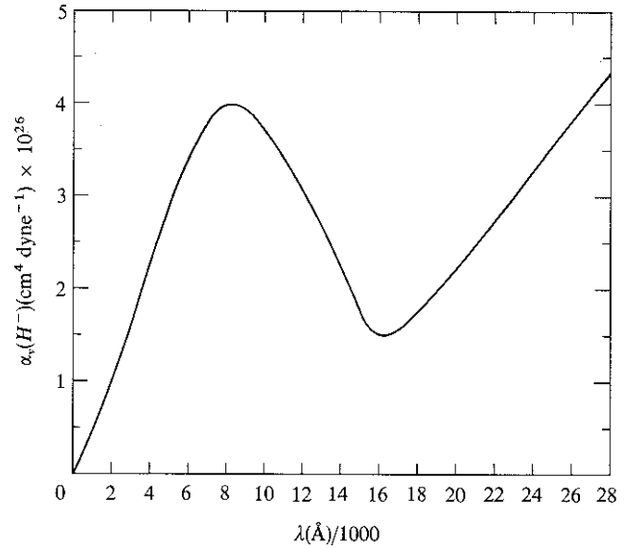



FIGURE 4-2
Bound–free and free–free opacity from $H^-$ at $T = 6300°$K. *Ordinate*: cross-section ($\times 10^{26}$) per neutral H atom and per unit electron pressure $p_e \equiv n_e kT$; *abscissa*: $\lambda/1000$ where $\lambda$ is in Å.

bound free near 15000 Å (1.5 $\mu$) and increases towards longer wavelengths. The summed absorption coefficient (see Figure 4-2) has a minimum at about 1.6 $\mu$; although other absorption processes act to wash out the minimum, the opacity for cool stars is smallest near this wavelength.

The determination of cross-sections for the two processes mentioned above is difficult and has been attempted both theoretically and experimentally. Very elaborate wave functions are required to give the desired accuracy. Pioneer calculations that gave fairly accurate values were carried out by Chandrasekhar and Breen (**162**). These were shown to be in accord with empirically deduced values for the absorption coefficient in the sun, and led to the firm identification of $H^-$ as the major opacity source in the solar atmosphere (see §3-6). More accurate values are now available for both the bound–free (**242**) and free–free (**604**) cross-sections; these are in good agreement with experimental values.

In LTE, the number of $H^-$ ions per cm$^3$ is given by a Saha formula [see equation (5-14)] that is of the form $n^*(H^-) = n(H)p_e\Phi(T)$ where $n(H)$ is the number of hydrogen atoms per cm$^3$, $p_e = n_e kT$, and $\Phi(T)$ contains the temperature dependence of the ionization equilibrium. The LTE opacity can thus be written $\kappa_\nu^*(H^-) = \alpha_\nu(H^-)n(H)p_e\Phi(T)(1 - e^{-h\nu/kT})$; departures from LTE may enter both in the calculation of $n(H^-)$ and in the stimulated emission

correction factor. Because $\kappa^*(H^-)$ is proportional to $p_e$, it is clear that it will be a more important opacity source in dwarfs than in giants. Also, because the electron density in G-type and cooler stars depends upon the abundance of the metals, $H^-$ will be a much weaker opacity source in Population II stars (which have low heavy-element abundances).

### OTHER IONS OF HYDROGEN

Hydrogen exists in two other forms that can contribute significantly to the opacity in stellar atmospheres, namely $H_2^+$ and $H_2^-$. The positive ion $H_2^+$ consists of a single electron shared by two protons; absorption cross-sections are given in (**72**) and (**117**). As the number density of $H_2^+$ is proportional to $n(H)n_p$, $H_2^+$ contributes significantly to the total opacity only for the temperature–pressure range where both neutral and ionized H-atoms exist simultaneously in appreciable numbers; i.e., where the hydrogen is about half-ionized. This range is characteristic of the A-stars, and $H_2^+$ makes about a 10 percent contribution to the opacity in the visible part of the spectrum (the $H_2^+$ absorption peak at $\lambda$1100 Å is swamped by the Balmer continuum of hydrogen).

The negative molecular ion $H_2^-$ exists only at relatively low temperatures, characteristic of the M-stars, and its free–free continuum makes a significant contribution at long wavelengths (the bound–free process is negligible). In this process, an electron passing near an $H_2$ molecule temporarily induces a dipole moment by polarization effects, and this moment can interact with the radiation field. The $H_2^-$ continuum tends to fill in the opacity minimum of $H^-$ at 1.6$\mu$. The free–free cross-section is given in (**592**).

### HELIUM

Helium is observed in stellar spectra in both its neutral and singly-ionized states. Because the ionization potential of neutral helium is 24.58 eV, it persists to temperatures characteristic of the B-stars, where hydrogen is already strongly ionized; in the O-stars, He II becomes a major opacity source. The threshold for abosprtion from the ground state of He I is at $\lambda$504 Å; the ultraviolet spectrum for $\lambda < 504$ Å is dominated by He I absorption for stars of types B0 and cooler. The excited states of helium fall into two groups, singlets and triplets, and each $(n, l, s)$ state has a different ionization energy. Roughly speaking, the ionization energies lie close to the hydrogenic value at the same $n$; thus helium contributes several absorption edges near each hydrogen edge (for $n \geqslant 2$). Because the excitation energy of even the lowest excited state is so large (19.72 eV), helium adds to the opacity in the visible regions of stellar spectra only in hot (B-type) stars. Generally, helium



is appreciably ionized before the excited states contribute to the opacity significantly. In a few stars the helium to hydrogen ratio is anomalous, and approaches or exceeds unity; here helium can dominate the opacity.

Helium is a three-body system, and exact wave functions cannot be obtained. A number of special methods can be applied to obtain accurate approximate wave functions [see (**87**, §§24–32; **577**, §§18.1 – 18.3)]; variational techniques applied to the ground state have been refined to the point of yielding very precise wave functions. The ground-state absorption coefficient calculated from an accurate Hartree–Fock wave function is given in (**603**); this agrees well with experimental values (**311**). Absorption cross-sections from the $2^3S, 2^3P, 2^1S, 2^1P$ levels have been calculated using accurate variational bound-state wave functions and close-coupling free-state wave functions (**332**). For higher excited states, precise cross-sections have not been published, and here one may use the quantum defect method.

Ionized helium is a hydrogenic ion with $Z = 2$. As energies in such ions scale as $Z^2$, the frequencies of the ionization edges, $\nu_n$, are larger by a factor of four, and the ground-state edge occurs at $\lambda$227 Å. This edge dominates the far ultraviolet spectrum of O-stars, except at the very highest temperatures where the helium becomes doubly ionized. The $n = 2$ edge of He II coincides with the hydrogen Lyman continuum; higher edges from states with *even* quantum numbers coincide with hydrogen edges from states with $n = n(\text{He II})/2$, while those from states with *odd* quantum numbers fall between the hydrogen edges.

Hydrogenic cross-sections can be used for He II, but the bound–free cross sections are a factor of $Z^4$ larger, and the free–free are a factor of $Z^2$ larger. The hydrogenic Gaunt factors apply if one evaluates them as functions of $(\nu/\nu_n)$. He II affects the visible spectrum only in stars of types B0 and hotter.

Finally, helium can give rise to a free–free opacity in cool stars. Cross-sections for this process are given in (**593**; **340**).

*Exercise 4-7*: Calculate the photoionization cross-sections of HeI from the four $n = 2$ states by the quantum defect method and compare these results with the more accurate values cited above.

## 4-5 Continuum Scattering Cross-Sections

As mentioned in Chapter 2, continuum radiation may be scattered as well as absorbed. In the latter case, photons are destroyed, and their energy contributed at least partially to the thermal content of the gas. In a scattering event, the photon is not destroyed, but merely redistributed in angle, and perhaps shifted slightly in frequency. Cross-sections for the two most important scattering processes in stellar atmospheres are derived in this section.

### THOMSON SCATTERING

The scattering of light by free electrons is referred to as *Thomson scattering*. The classical formula for this process can be obtained directly from equation (4-32) by noting that for an unbound electron both the resonant frequency $\omega_0$ and the damping parameter are zero. Thus we find

$$\sigma_e = (8\pi e^4/3m^2c^4) = 6.65 \times 10^{-25} \text{ cm}^2 \qquad (4\text{-}127)$$

Note that this cross-section is *independent of frequency*. The Thomson cross-section has been verified by quantum mechanical calculations in the limit of low photon energies; i.e., $h\nu \ll mc^2$. At high photon energies ($\lambda <$ 1 Å, in the $X$-ray region) one must employ the *Klein–Nishina* formula (**293**, §22; **392**, 433), which predicts a smaller cross-section; in practice the departure of $\sigma$ from $\sigma_e$ can be ignored in stellar atmospheres work (except for X-ray binaries).

In the derivation of equation (4-32) an angle averaging was performed, and thus the angular dependence of the scattering coefficient was suppressed; the correct angular dependence is given by the dipole phase function in equation (2-19). In stellar atmospheres applications, this angular dependence can almost always be ignored and the process considered to be isotropic. Frequency redistribution in the laboratory frame caused by Doppler shifts from the electrons' motions have also been neglected; these will be considered in Chapter 13 (see Exercises 13-5 and 13-6). In the continuum, the frequency redistribution just mentioned can be ignored; near a spectral line, it may be necessary to take it into account. Electron scattering is one of the most important opacity sources in hot stars (e.g., the O-stars).

### RAYLEIGH SCATTERING

The term *Rayleigh scattering* refers to the scattering of radiation by bound systems, such as atoms or molecules, at frequencies much lower than characteristic transition frequencies of the system. Again using equation (4-32), the process can be described by representing real transitions of the scatterer with equivalent classical oscillators of appropriate strengths $f_{ij}$, and resonant frequencies $\omega_{ij}$ equal to the actual transition frequencies. Then for $\omega \ll \omega_{ij}$, equation (4-32) simplifies to yield

$$\sigma(\omega) = (8\pi e^4/3m^2c^4) f_{ij}\omega^4/(\omega_{ij}^2 - \omega^2)^2 = \sigma_e f_{ij}\omega^4/(\omega_{ij}^2 - \omega^2)^2 \qquad (4\text{-}128)$$

Far from the resonant frequency, $\sigma(\omega)$ varies as $\omega^4$ or $\lambda^{-4}$, which leads to a strong color dependence of the scattered radiation; a well-known example of this dependence is the blue color of the sky, resulting from sunlight scattered by molecules of air.



Rayleigh scattering can be important in the atmospheres of stars of moderate temperature (spectral types G and K). Here most of the hydrogen is neutral and in the ground state. The resonant frequencies corresponding to the Lyman lines ($1 \rightarrow n$) lie far in the ultraviolet, and visible photons interact with these transitions by the Rayleigh scattering mechanism. The macroscopic scattering coefficient is obtained by summing over all lines and multiplying by the density of hydrogen in the ground state. Rayleigh scattering from neutral hydrogen can dominate the total opacity at relatively low temperatures and high frequencies; see the graphs in (**97**) and (**651**). Moreover, in stars with low metal abundances (Population II stars), the number of free electrons (coming mainly from the metals) is greatly reduced; accordingly the opacity from $H^-$ is diminished, and thus the importance of Rayleigh scattering is much enhanced. Molecular hydrogen, $H_2$, may also scatter radiation in an analogous fashion. The cross-section per molecule (**189**) is comparable to the cross-section per atom for atomic hydrogen. At low temperatures (e.g., in M-stars), $H_2$ is much more prevalent than atomic H, and thus molecular Rayleigh scattering dominates.

It should be noted that in a continuum scattering process there is no analogue of the stimulated emission that occurs in absorption processes. Thus in the macroscopic scattering coefficient there is no correction factor for stimulated emission such as appears in equations (4-98) and (4-99).

# 5

# *The Equations of Statistical Equilibrium*

Stellar atmospheres are regions of high temperature and low density. Therefore the gas consists mainly of single atoms, ions, and free electrons; in cooler stars molecules also form. Because of the low densities, the material always behaves as a perfect gas. The *state* of the gas is specified when we know the distribution of the particles over all available bound and free energy levels—i.e., when we know the *occupation numbers* of these levels. We then have the information required to compute the gas pressure, mass density, opacity, emissivity, etc. of the material.

To specify occupation numbers, we must deal with the phenomena of excitation and ionization of each chemical species in the gas. One approach is to *assume* that we may apply the equilibrium relations of statistical mechanics and thermodynamics at local values of the temperature and density; this is the *local thermodynamic equilibrium* (LTE) approach. As we shall see, LTE provides an extremely convenient method for computing the particle distribution functions. One of the fundamental properties of stellar atmospheres, however, is the presence of an intense radiation field whose character is very different from the equilibrium Planck distribution. This radiation field interacts strongly with the material via radiative excitations and photo-



ionizations (and their inverses), and thus acts to help determine the occupation numbers of the gas; we shall show in fact, that *radiative transitions dominate* the state of the gas. In this case the occupation numbers must be determined from *equations of statistical equilibrium*, which specify all of the microprocesses that produce transitions from one atomic state to another.

The fact that the state of the material depends upon the radiation field introduces the essential difficulty of stellar atmospheres theory for, as we mentioned in Chapter 2, the radiation field, in turn, depends on the occupation numbers via the absorptivity and emissivity and their effects upon the transfer of radiation through the atmosphere. Thus what is required is a completely *self-consistent simultaneous solution of both the radiative transfer and statistical equilibrium equations.* This is a difficult problem in general, and its solution occupies the bulk of Chapters 7, 11, and 12 of this book. For the present we shall only show that there are strong expectations that the state of the material will depart from that predicted by LTE, which is therefore at best a computational expedient. If in any particular case the occupation numbers obtained from the general analysis happen to agree with those predicted by LTE, then one may legitimately use the LTE assumption; but for a wide range of problems (line-formation in particular), such agreement is not generally attained (nor can we accurately predict a priori when it will be for most cases of interest!).

## 5-1 Local Thermodynamic Equilibrium

In thermodynamic equilibrium, the state of the gas (i.e., the distribution of atoms over bound and free states) is specified uniquely by two thermodynamic variables (we shall choose the absolute temperature $T$ and the total particle density $N$) via the well-known equilibrium relations of statistical mechanics. These relations will not all be derived in this chapter, as they are easily found in standard texts [see, e.g., (**565**, Chaps. 12, 14, and 15; **11**, Chap. 3)] but will be summarized in forms useful for further developments in this book. The assumption of LTE asserts that we may employ these same relations in a stellar atmosphere at the *local* values $T(\mathbf{r})$ and $N(\mathbf{r})$ despite the gradients that exist in the atmosphere. This simple assumption is actually a very strong one, for it implies that we propose to calculate the above-mentioned distribution functions without reference to the physical ensemble in which the given element of material is found. Thus it is assumed that it is irrelevant whether the material is contained within an equilibrium cavity (the classical *hohlraum*), an atmosphere with a strong radiation field, or in the exhaust of a space vehicle, despite the obvious dissimilarities of these situations! In LTE, we have a purely *local* theory, which makes no allowance for coupling of the state of one element of gas with that of another,

say by radiative exchange (except as may be imposed by certain global constraints on the atmosphere—e.g., hydrostatic or radiative equilibrium). Moreover, in LTE the absolute temperature $T$ has a quite general significance. The *same* $T$ applies in the calculation of the velocity distribution functions of atoms, ions, and electrons; the distribution of atoms and ions over all states (Boltzmann–Saha equations); and the distribution of thermal emission (Planck function). In short, the full implications of the LTE assumption are quite sweeping. It is this very fact which makes it so effective in reducing the complexity of the equations, and at the same time so difficult to justify physically and so vulnerable to error.

### THE MAXWELLIAN VELOCITY DISTRIBUTION

The probability, in thermodynamic equilibrium, that a particle of mass $m$ at temperature $T$ has a velocity on the range $(\mathbf{v}, \mathbf{v} + \mathbf{dv})$ is given by the *Maxwellian velocity distribution*

$$f(\mathbf{v})\, dv_x\, dv_y\, dv_z = \left(\frac{m}{2\pi kT}\right)^{\frac{3}{2}} \exp\left[-m(v_x^2 + v_y^2 + v_z^2)/2kT\right] dv_x\, dv_y\, dv_z \quad (5\text{-}1)$$

or, in terms of speeds on the range $(v, v + dv)$

$$f(v)\, dv = \left(\frac{m}{2\pi kT}\right)^{\frac{3}{2}} \exp(-mv^2/2kT) 4\pi v^2\, dv \quad (5\text{-}2)$$

These distributions may be characterized in terms of the *most probable speed*

$$v_0 = (2kT/m)^{\frac{1}{2}} = 12.85(T/10^4 A)^{\frac{1}{2}} \text{ km/sec} \quad (5\text{-}3)$$

where $A$ is the atomic weight of the particle. Related parameters are the root-mean-square speed $\langle v^2 \rangle^{\frac{1}{2}} = (3kT/m)^{\frac{1}{2}}$, and the root-mean-square velocity in one component (e.g., along the line of sight) $\langle v_x^2 \rangle^{\frac{1}{2}} = (kT/m)^{\frac{1}{2}}$.

### THE BOLTZMANN EXCITATION EQUATION

In thermodynamic equilibrium at temperature $T$, atoms are distributed over their bound levels according to the *Boltzmann excitation equation.* Let $n_{ijk}$ denote the number density of atoms in excited level $i$ of ionization state $j$ of chemical species $k$. Let $j = 0$ denote neutral atoms, $j = 1$ singly ionized atoms, etc. Measure the excitation energy $\chi_{ijk}$ relative to the ground state of the atom. Let $g_{ijk}$ denote the statistical weight assigned to the level to account for degenerate sublevels (e.g., the $2J + 1$ $m$-states in the absence of a magnetic field). Then, according to the Boltzmann law the population of



any excited level is

$$(n_{ijk}/n_{0jk})^* = (g_{ijk}/g_{0jk}) \exp(-\chi_{ijk}/kT) \quad (5\text{-}4)$$

where the subscript 0 denotes the ground level and the superscript * denotes LTE. For any two excited levels $l$ and $m$,

$$\begin{aligned}(n_{mjk}/n_{ljk})^* &= (g_{mjk}/g_{ljk}) \exp\left[-(\chi_{mjk} - \chi_{ljk})/kT\right] \\ &= (g_{mjk}/g_{ljk}) \exp(-h\nu_{lm}/kT) \end{aligned} \quad (5\text{-}5)$$

where $h\nu_{lm}$ is the energy of a photon that equals the energy difference between the levels. In calculations of ionization equilibria, we typically wish to know the *total* number of atoms in a particular ionization state, which can be written as

$$\begin{aligned} N_{jk}^* &= \sum_i n_{ijk}^* = (n_{0jk}^*/g_{0jk}) \sum_i g_{ijk} \exp(-\chi_{ijk}/kT) \\ &= (n_{0jk}^*/g_{0jk}) U_{jk}(T) \end{aligned} \quad (5\text{-}6)$$

where

$$U_{jk}(T) \equiv \sum_i g_{ijk} \exp(-\chi_{ijk}/kT) \quad (5\text{-}7)$$

is called the *partition function.* A form of equation (5-4) customarily used in classical curve-of-growth analyses of spectra (see §§10-3 and 10-4) is

$$(n_{ijk}/N_{jk})^* = g_{ijk} \exp(-\chi_{ijk}/kT)/U_{jk}(T) \quad (5\text{-}8)$$

The partition function is tedious to compute, and for some atoms and ions (e.g., of the rare earths) our knowledge of the term structure is so incomplete that we lack the needed data. Tabulations of partition functions (**11**, 115–117) and convenient analytical fits with approximation formulae (**103**; **220**) are available. Often a fair estimate is just $g_{0jk}$ or a sum over a few low-lying states. Note that formally the partition function in equation (5-7) diverges if the sum extends over all (an infinite number) of bound states, for a lower bound to each term of the sum is $\exp(-\chi_{Ijk}/kT)$ (where $\chi_{Ijk}$ is the ionization potential), which is nonzero. This problem is not a physical one, for in reality the highest levels cannot remain bound because they are strongly perturbed by neighboring atoms and ions. To estimate the effect we might suppose that the only levels that are bound are those contained within the average volume available to ions. For a particle density $N$, the mean interatomic distance is $r_0 = (3/4\pi N)^{\frac{1}{3}}$, and for a hydrogenic ion of charge $Z$ the radius of the state of principal quantum number $n$ is $r_n = n^2 a_0/Z$ where $a_0$ is the Bohr radius $= 5.3 \times 10^{-9}$ cm. If we set $r_n = r_0$ and choose

a typical value of $N \sim 10^{15}$ we find $n \approx 30Z^{\frac{1}{2}}$, so clearly the sum is finite. A more accurate calculation (see §9-4) shows that in a plasma with charge density $n_e$ and temperature $T$, an ion of charge $Z$ suffers a lowering of the ionization potential by an amount $\Delta\chi \approx 3 \times 10^{-8}\, Z n_e^{\frac{1}{2}} T^{-\frac{1}{2}}$ eV. For hydrogenic energy levels this implies $n_{\max}^2 \approx 4 \times 10^8\, Z n_e^{-\frac{1}{2}} T^{\frac{1}{2}}$ or $n_{\max} \sim 60Z^{\frac{1}{2}}$ for $n_e \sim 10^{14}$ and $T \sim 10^4$.

### THE SAHA IONIZATION EQUATION

Above the discrete bound eigenstates of an atom there exists a *continuum* of levels in which the electron is unbound and has a nonzero kinetic energy. The energy above the ground level at which this continuum begins is called the *ionization potential* $\chi_I$. The relative numbers of atoms and ions in successive stages of ionization can be computed from the *Saha ionization equation*, which we shall derive as an extension of the Boltzmann formula to free states.

Consider a process in which an atom of species $k$ is ionized from its ground level, resulting in an ion in the ground level plus a free electron in the continuum moving with speed $v$. The energy required to carry out this process is $\chi_{I,0,k} + \frac{1}{2}mv^2$ (where we have used notation parallel to that used in our discussion of the Boltzmann formula). The statistical weight of the initial state is $g_{0,0,k}$, and the statistical weight of the final state (ion + electron) may be written $g(v) = g_{0,1,k} \times g_{\text{electron}}$. If we use $n_{0,1,k}(v)$ to denote the number of ions in the ground level with accompanying free electron with speed in the range $(v, v + dv)$, we may apply equation (5-4) to write

$$[n_{0,1,k}(v)/n_{0,0,k}]^* = [g(v)/g_{0,0,k}] \exp[-(\chi_{I,0,k} + \tfrac{1}{2}mv^2)/kT] \tag{5-9}$$

We identify $g_{\text{electron}}$ with the number of phase space elements available to the free electron, which, according to quantum statistics, is

$$g_{\text{electron}} = 2(dx\, dy\, dz\, dp_x\, dp_y\, dp_z)/h^3$$

where the factor of 2 accounts for the two possible orientations of the electron spin. We choose the space-volume element to contain exactly one free electron, and make the substitution $dx\, dy\, dz = n_e^{-1}$. We rewrite the momentum volume element in terms of the electron's speed,

$$dp_x\, dp_y\, dp_z = 4\pi p^2\, dp = 4\pi m^3 v^2\, dv$$

Then equation (5-9) becomes

$$[n_{0,1,k}(v)/n_{0,0,k}]^* = 8\pi m^3 h^{-3} (g_{0,1,k}/g_{0,0,k}) n_e^{-1} \exp[-(\chi_{I,0,k} + \tfrac{1}{2}mv^2)/kT]\, v^2\, dv \tag{5-10}$$



Now, summing over all final states by integrating over the electron velocity distribution we obtain

$$(n_{0,1,k} n_e / n_{0,0,k})^* = 8\pi m^3 h^{-3} (g_{0,1,k}/g_{0,0,k}) (2kT/m)^{\frac{3}{2}} \exp(-\chi_{I,0,k}/kT) \int_0^\infty e^{-x^2} x^2\, dx \tag{5-11}$$

or, evaluating the integral,

$$n^*_{0,0,k} = n^*_{0,1,k} n_e^* \tfrac{1}{2} (h^2/2\pi mkT)^{\frac{3}{2}} (g_{0,0,k}/g_{0,1,k}) \exp(\chi_{I,0,k}/kT) \tag{5-12}$$

which is a basic form of *Saha's equation*. Note that in the derivation we made no explicit reference to the ionization state of the initial "atom," hence we may extend equation (5-12) to apply between any two successive stages of ionization

$$n^*_{0jk} = n^*_{0,j+1,k} n_e^* \tfrac{1}{2} (h^2/2\pi mkT)^{\frac{3}{2}} (g_{0jk}/g_{0,j+1,k}) \exp(\chi_{Ijk}/kT) \tag{5-13}$$

If, further, we apply Boltzmann's formula, equation (5-4), we obtain an expression for the occupation number of any state of ion $j$ in terms of the temperature, electron density, and ground state population of ion $j + 1$, namely

$$\begin{aligned} n^*_{ijk} &= n_{0,j+1,k} n_e (g_{ijk}/g_{0,j+1,k}) C_I T^{-\frac{3}{2}} \exp[(\chi_{Ijk} - \chi_{ijk})/kT] \\ &\equiv n_{0,j+1,k} n_e \Phi_{ijk}(T) \end{aligned} \tag{5-14}$$

Equation (5-14) is the most useful form of Saha's equation for the formalism we shall employ, and will be used to *define* LTE populations in the full non-LTE equations of statistical equilibrium (for this reason the superscript * on $n_{0,j+1,k}$ and $n_e$ has been omitted). The constant has the value $C_I = 2.07 \times 10^{-16}$ in cgs units.

By applying equation (5-6), we may rewrite equation (5-14) as

$$n^*_{ijk} = N^*_{j+1,k} n_e [g_{0,j+1,k}/U_{j+1,k}(T)] \Phi_{ijk}(T) \equiv N^*_{j+1,k} n_e \tilde{\Phi}_{ijk}(T) \tag{5-15}$$

Further, by summing over all bound levels of the lower ionization stage and again using equation (5-6), we obtain an equation for the ratio of the total number of atoms in successive stages of ionization:

$$(N_{jk}/N_{j+1,k})^* = n_e [U_{jk}(T)/U_{j+1,k}(T)] C_I T^{-\frac{3}{2}} \exp(\chi_{Ijk}/kT) \equiv n_e \tilde{\Phi}_{jk}(T) \tag{5-16}$$

By recursive application of equation (5-15) between successive stages of ionization, we can obtain an expression for the fraction of atoms of chemical

species $k$ in ionization stage $j$ relative to the total number of atoms of that species:

$$\begin{aligned} f_{jk}(n_e, T) &\equiv (N_{jk}/N_k)^* \\ &= \frac{(N_{J-1,k}/N_{Jk})^* \cdots (N_{jk}/N_{j+1,k})^*}{1 + (N_{J-1,k}/N_{Jk})^* + (N_{J-1,k}/N_{Jk})^*(N_{J-2,k}/N_{J-1,k})^* + \cdots + (N_{J-1,k}/N_{Jk})^* \cdots (N_{0k}/N_{1k})^*} \\ &= \prod_{l=j}^{J_k-1} [n_e \tilde{\Phi}_{lk}(T)] \Bigg/ \sum_{m=0}^{J_k} \prod_{l=m}^{J_k-1} [n_e \tilde{\Phi}_{lk}(T)] \\ &\equiv P_{jk}(n_e, T)/S_k(n_e, T), \quad (j = 1, \ldots, J_k) \end{aligned} \tag{5-17}$$

where $J_k$ is the last ionization stage of species $k$ considered. We observe the convention that the product term for $l = J_k$ (which formally becomes void) is replaced by unity in both the numerator and denominator.

Consideration of the above results shows that, *if we know* $(n_e, T)$, then we may determine, for any chemical species $k$, the fraction in any chosen ionization stage from equation (5-17), and in any particular excitation state from equation (5-15). If, in addition, we know the total number density of atoms of this species, we can obtain absolute occupation numbers $n_{ijk}$. In practice this procedure is useful in LTE calculations of line spectra where we are given a model atmosphere that specifies $n_e(z)$ and $T(z)$. In computation of the model itself, however, we generally do not know $n_e(z)$, but rather the total particle density $N(z)$; we must then determine $n_e$, and as can be seen from equation (5-17), this implies we must solve a nonlinear set of equations. Let us therefore now consider methods of solving the nonlinear problem.

## 5-2 The LTE Equation of State for Ionizing Material

The Saha–Boltzmann relations allow a computation of the fraction of each chemical species in various stages of ionization, and the number of free electrons that each contributes to the plasma. Stellar atmospheres consist of a mixture of elements with widely differing ionization potentials; in general some of the species may be neutral while others are singly or multiply ionized. Usually the transition from one ionic stage to the next occurs fairly abruptly with increasing temperature, and normally a particular chemical species exists essentially entirely in two successive ionization stages. This provides a sensitive diagnostic tool to infer the temperature structure of a stellar atmosphere, for it implies that ratios of line strengths of two successive ionic spectra (e.g., He I and He II, or Ca I and Ca II) will vary



rapidly as a function of temperature. In fact, this was the basis upon which the first understanding of the spectral sequence as a temperature sequence was built by Saha (**546**; **547**), Pannekoek (**495**), Cecilia Payne (**501**), and Fowler and Milne (**222**; **223**). In normal stellar atmospheres hydrogen is by far the most abundant constituent, and helium is next most abundant with $N(\text{He})/N(\text{H}) \approx 0.1$. The heavier elements have much smaller abundances relative to hydrogen [see, e.g., (**252**) for element abundances in the solar atmosphere]. At typical temperatures in the solar atmosphere (6000°K) hydrogen is essentially neutral, and the electrons are contributed mainly by the "metals" such as Na, Mg, Al, Si, Ca, and Fe. At higher temperatures, characteristic of the A-stars (10,000°K), hydrogen ionizes and becomes the dominant source of electrons. At very high temperatures, characteristic of the O- and early B-stars, helium ionizes and makes an appreciable contribution of electrons.

### CHARGE AND PARTICLE CONSERVATION

In calculations of stellar atmospheres we specify the gas pressure from the equation of hydrostatic equilibrium. Thus, given $p_g$ and $T$, we know the total number density $N$ from the relation

$$p_g = NkT = (N_{\text{atoms}} + N_{\text{ions}} + n_e)kT = (N_{\text{N}} + n_e)kT \tag{5-18}$$

Here $N_{\text{N}}$ denotes the density of "nuclei"; i.e., atoms and ions of all types. In equation (5-18) and subsequent equations of this section we suppress the "*" that denotes LTE for notational simplicity. We define the *abundance* $\alpha_k$ of chemical species $k$ to be such that $N_k = \alpha_k N_{\text{N}}$ where $\Sigma_k \alpha_k \equiv 1$. Then

$$N_k = \alpha_k (N - n_e) \tag{5-19}$$

summarises the constraint of *particle conservation* (i.e., $\Sigma N_k = N_{\text{N}}$). In addition we require the plasma to be electrically neutral; then the number of free electrons equals the total ionic charge, and the condition of *charge conservation* reads

$$n_e = \sum_k \sum_{j=1}^{J_k} j N_{jk} = \sum_k N_k \sum_{j=1}^{J_k} j f_{jk}(n_e, T) = (N - n_e) \sum_k \alpha_k \sum_{j=1}^{J_k} j f_{jk}(n_e, T) \tag{5-20}$$

As mentioned above, if we know $(n_e, T)$ we may calculate $N$ and the $f_{jk}$ directly. But if we know $(N, T)$, we must find $n_e$ from a nonlinear equation. Before the availability of electronic computers this problem was solved by constructing tables of $\log p_g(T, \log p_e)$ (here $p_e \equiv n_e kT$), in which interpolations could be made to find $\log p_e(T, \log p_g)$. Examples of such tables are

given in (**11**, 130) and (**638**, 104). We shall develop a different procedure, along lines suggested by L. H. Auer, that is better suited for machine computation, and that fits into the overall approach of Chapter 7 for the computation of model atmospheres. But first consider an instructive example that yields physical insight in limiting cases.

Suppose that the gas consists only of hydrogen ($\chi_H = 13.6$ eV) and one metal "M" with a single ionization stage of much lower ionization potential (say $\chi_M = 4$ or 5 eV) and an abundance relative to hydrogen $\alpha_M \ll 1$. At high temperatures where the hydrogen is appreciably ionized, it will contribute most of the electrons; at lower temperatures the hydrogen is neutral, and $n_e$ is determined by $f_M$, the ionization fraction of the metal. The number of particles of all types is

$$N = n_H(1 + f_H) + \alpha_M n_H(1 + f_M) \tag{5-21}$$

while the number of electrons is

$$n_e = n_p + n_{M+} = n_H(f_H + \alpha_M f_M) \tag{5-22}$$

Then
$$p_e/p_g = (f_H + \alpha_M f_M)/[1 + f_H + \alpha_M(1 + f_M)] \tag{5-23}$$

At high enough temperatures, $f_H \to 1$, and as $\alpha_M \ll 1$, $(p_e/p_g) \to \frac{1}{2}$. At intermediate temperatures, where $\alpha_M \ll f_H \ll 1$, and at the same time $f_M \approx 1$, $(p_e/p_g) \approx f_H$. At low temperatures, $f_H \to 0$ while $f_M/f_H \gg 1$, hence $(p_e/p_g) \to \alpha_M f_M$. We thus see that at high temperatures the metals are essentially irrelevant to the determination of $p_e/p_g$, while at low temperatures they play a crucial role. In particular, note that the metal abundance enters directly in fixing $p_e$ in cool stars; this is important because the dominant opacity source in cooler atmospheres is absorption by the $H^-$ ion, and $n(H^-)/n(H)$ is proportional to $n_e$. Thus in these stars the metal abundance fixes the opacity as well.

For a *pure hydrogen gas*, equations (5-16), (5-19), and (5-20) may be solved analytically to obtain

$$n_e(H) = \tilde{\Phi}_H^{-1}[(N\tilde{\Phi}_H + 1)^{\frac{1}{2}} - 1] \tag{5-24}$$

which shows that at low degrees of ionization, $n_e \sim N^{\frac{1}{2}}$ for a given $T$.

*Exercise 5-1:* Derive equation (5-24).

If only the metal in our two-component gas described above is ionized ($f_H \ll \alpha_M$) then we have

$$n_e(M) \approx \tilde{\Phi}_M^{-1}\left\{\left[\alpha_M N\tilde{\Phi}_M + \frac{1}{4}(1 + 2\alpha_M)\right]^{\frac{1}{2}} - \frac{1}{2}(1 + \alpha_M)\right\} \tag{5-25}$$



A fairly good estimate of $n_e$ can be obtained from equations (5-24) and (5-25) if $\alpha_M \ll 1$ and $\chi_M \ll \chi_H$ by writing $n_e \approx n_e(H) + n_e(M)$.

### SOLUTION BY LINEARIZATION

Let us now turn to the problem of determining $n_e$ for a given value of $(N, T)$ by means of an iterative linearization procedure (generalized Newton–Raphson method). We shall describe the procedure in fair detail because it is a simple example of the approach we shall use in more complicated cases (e.g., the non-LTE rate equations and the transfer equation). The only equation to be solved (contrast this with the non-LTE case, cf. §5.5!) is equation (5-20) where $f_{jk}(n_e, T)$ is given by equation (5-17). Suppose that we have an initial estimate of the electron density, $n_e^0$; suppose also we find that using $n_e^0$ to evaluate the righthand side of equation (5-20) yields a density $n_e^1 \neq n_e^0$. It is then clear that the true density differs from $n_e^0$, so we write $n_e = n_e^0 + \delta n_e$ where $\delta n_e$ is to be determined in such a way as to satisfy equation (5-20) exactly. Because the equation is nonlinear, we cannot determine this $\delta n_e$ exactly, but on the supposition that $\delta n_e/n_e^0 \ll 1$, we can *estimate* $\delta n_e$ by expanding all terms to first order and solving for $\delta n_e$. Then we have

$$n_e^0 + \delta n_e \approx [(N - n_e^0 - \delta n_e)\cdot\tilde{\Sigma}(n_e^0, T)] + (N - n_e^0)[\partial\tilde{\Sigma}(n_e, T)/\partial n_e]_{n_e^0}\,\delta n_e \tag{5-26}$$

or
$$\delta n_e \approx [(N - n_e^0)\tilde{\Sigma} - n_e^0][1 + \tilde{\Sigma} - (N - n_e^0)(\partial\tilde{\Sigma}/\partial n_e)]^{-1} \tag{5-27}$$

where
$$\tilde{\Sigma}(n_e, T) \equiv \sum_k \alpha_k S_k^{-1}(n_e, T) \sum_{j=1}^{J_k} jP_{jk}(n_e, T) \tag{5-28}$$

Note that we may rewrite the functions $P(n_e, T)$ and $S(n_e, T)$ as

$$P_{jk}(n_e, T) \equiv \prod_{i=j}^{J_k - 1} [n_e\tilde{\Phi}_{ik}(T)] = n_e^{(J_k - j)} \prod_{jk}(T) \tag{5-29}$$

and
$$S_k(n_e, T) \equiv \sum_{j=0}^{J_k} P_{jk}(n_e, T) = \sum_{j=0}^{J_k} n_e^{(J_k - j)} \prod_{jk}(T) \tag{5-30}$$

The value of $\delta n_e$ given by equation (5-27) will not be exact, so we *iterate* the procedure by using a *new* estimate $n_e^0(\text{new}) = n_e^0(\text{old}) + \delta n_e$ to re-evaluate $\tilde{\Sigma}$ and $\partial\tilde{\Sigma}/\partial n_e$, and to compute yet another value of $\delta n_e$.

The convergence of this procedure is *quadratic* (if our original estimate lies within the range of convergence) so, if the first fractional error $\delta n_e/n_e$ is $\varepsilon$, subsequent iterations will produce corrections of order $\varepsilon^2$, $\varepsilon^4$, $\varepsilon^8$, etc.,

which implies that one can obtain the result to the desired accuracy quickly. It is also worth noting that the derivative $\partial\tilde{\Sigma}/\partial n_e$ can be evaluated *analytically*:

$$(\partial\tilde{\Sigma}/\partial n_e) = \sum_k \alpha_k \left[ S_k^{-1} \sum_j j(\partial P_{jk}/\partial n_e) - S_k^{-2}(\partial S_k/\partial n_e) \sum_j j P_{jk} \right] \quad (5\text{-}31)$$

where $(\partial P_{jk}/\partial n_e)$ and $(\partial S_k/\partial n_e)$ follow immediately from equations (5-29) and (5-30) and produce a compact expression for (5-31). In general, the derivatives appearing in linearization procedures can be estimated numerically; however, we shall usually be able to obtain analytical derivatives, and experience has shown that in this way we obtain better control of the calculation.

Finally, having obtained a satisfactory value for $n_e$, and, as a byproduct the $f_{jk}$, we may calculate any particular occupation number from equation (5-15)

$$n_{ijk} = N_{j+1,k} n_e \tilde{\Phi}_{ijk}(T) = \alpha_k (N - n_e) n_e f_{j+1,k}(n_e, T) \tilde{\Phi}_{ijk}(T) \quad (5\text{-}32)$$

this completes the computation of the LTE equation of state.

The procedure outlined above has a larger significance than indicated thus far. We have assumed that $N$ and $T$ are given. But these quantities follow from constraints of pressure and energy balance, and in general are known only approximately at any particular stage of calculation of a model. As we shall see in Chapter 7, we may apply the linearization procedure to *all* the variables involved, and hence we shall need to evaluate the response of the occupation numbers to the perturbations $\delta N$ and $\delta T$. Perturbing equation (5-20) we obtain

$$\begin{aligned} n_e + \delta n_e = {} & (N + \delta N - n_e - \delta n_e)\tilde{\Sigma} \\ & + (N - n_e)[(\partial\tilde{\Sigma}/\partial n_e)\,\delta n_e + (\partial\tilde{\Sigma}/\partial T)\,\delta T] \end{aligned} \quad (5\text{-}33)$$

or, assuming that $n_e$ *is* a solution of equation (5-20) at the current values of $(N, T)$,

$$\begin{aligned} \delta n_e &= [1 + \tilde{\Sigma} - (N - n_e)(\partial\tilde{\Sigma}/\partial n_e)]^{-1}[\tilde{\Sigma}\,\delta N + (N - n_e)(\partial\tilde{\Sigma}/\partial T)\delta T] \\ &\equiv (\partial n_e/\partial N)_T\,\delta N + (\partial n_e/\partial T)_N\,\delta T \end{aligned} \quad (5\text{-}34)$$

where again $\partial\tilde{\Sigma}/\partial T$ may be evaluated analytically. Further, from equation (5-32) we may develop an expression for $\delta n_{ijk}$ of the form $\delta n_{ijk} = \Delta_1\,\delta N + \Delta_2\,\delta T + \Delta_3\,\delta n_e$, which can be collapsed down by use of equation (5-34) to an expression of the form

$$\delta n_{ijk} = (\partial n_{ijk}/\partial N)_T\,\delta N + (\partial n_{ijk}/\partial T)_N\,\delta T \quad (5\text{-}35)$$



*Exercise 5-2*: Obtain expressions for the coefficients in equation (5-35) in terms of $P_{jk}$, $S_k$, and $\tilde{\Phi}_{ijk}$, and their derivatives.

Equations (5-34) and (5-35) provide the information we shall require in §7-2 to find the response of the opacity and emissivity $(\delta\chi, \delta\eta)$ to changes in the model structure $(\delta N, \delta T)$.

*Exercise 5-3*: Show that $\delta N_{Jk}$, for the *last* ionization stage of element $k$, has a particularly simple form because $f_{Jk}$ involves only $S_k$. Then show that expressions for $\delta N_{jk}$ of lower ions can be evaluated *recursively* from equation (5-16), and that these lead from equation (5-15) to simple expressions of the form of equation (5-35) for $\delta n_{ijk}$.

## 5-3 The Microscopic Requirements of LTE

Before we develop the equations of statistical equilibrium, it is worthwhile to discuss qualitatively the microscopic requirements of LTE. An interesting commentary on these requirements by K. H. Böhm may be found in (**261**, Chap. 3); we shall summarize and discuss this analysis here along with other material of relevance.

### DETAILED BALANCE

In thermodynamic equilibrium, the rate at which each process occurs is exactly balanced by the rate at which its inverse occurs, for *all* processes; i.e., each process is in *detailed balance*. This is a very strong requirement, and it proves to be very useful in constructing relations among rate coefficients (recall the use of this procedure in Chapter 4). We may classify processes that produce transitions from one state to another (bound or free) into two broad categories: radiative and collisional. Collisional processes are the processes invoked in statistical mechanics to establish equilibrium, and can be expected to be in detailed balance *whenever the velocity distribution of the colliding particles is the equilibrium* (i.e., Maxwellian) *distribution*. We shall show below that this can be expected to be the case in stellar atmospheres. Furthermore, we may make the same statement about processes which are essentially collisional in character, even though a photon is emitted (e.g., free–bound radiative recombination and free–free emission); we can therefore use detailed balancing arguments to calculate the rates of these processes when convenient to do so. In contrast, radiative processes (e.g., photoexcitation, photoionization) depend directly upon the character of the radiation field, and will be in detailed balance *only* if the radiation field is isotropic and has a Planck distribution. We shall show below that this is *not* the case in stellar atmospheres.

If some processes are in detailed balance while others are not, the final occupation numbers will be determined by a competition among them and may depart more or less strongly from an equilibrium distribution. LTE will be valid in the *deepest* layers of stellar atmospheres where densities are high and the collision rates become large, and the optical depth is so large that no photon escapes from the atmosphere before being thermalized, so that the radiation field approaches the Planck function. But in the observable layers, precisely the opposite regime is found.

### THE NATURE OF THE RADIATION FIELD

A stellar atmosphere is not in any sense a closed system in equilibrium at a uniform temperature. Indeed the opposite situation prevails: radiation flows freely from the surface layers of the star into essentially empty space, which implies that the radiation field is decidedly anisotropic, and that the atmosphere has a large temperature gradient. The radiation field at any point is the integrated result of emissions and absorptions over the entire (possibly *large*) volume within which a photon can travel from its point of emission to the test point. This volume may include the boundary surface and empty space beyond, with a consequent reduction of intensity, as well as layers of higher temperatures and densities from which intense radiation originates. The radiation field therefore is distinctly nonlocal in nature, and has an absolute intensity, directional distribution, and frequency spectrum that may have no resemblance whatever to the local equilibrium distribution $B_\nu(T)$. Radiative rates may therefore be far from their equilibrium values, and thus tend to drive the material away from LTE.

The radiation field is plainly anisotropic because the radiating surface subtends a solid angle less than $4\pi$, and essentially no radiation enters from the surrounding void. We may describe this geometrical effect by introducing a *dilution factor* $W$ defined to be $\omega_*/4\pi$ where $\omega_*$ is the solid angle subtended by the stellar disk.

*Exercise 5-4*: Show that

$$W = \frac{1}{2}\{1 - [1 - (r_*/r)^2]^{\frac{1}{2}}\} \tag{5-36}$$

where $r_*$ is the radius of the radiating surface and $r$ denotes the position of the observer. Show that for $r_*/r \ll 1$, $W = \frac{1}{4}(r_*/r)^2$

As defined, $W$ clearly measures the factor by which the energy density in the radiation field is reduced as the source of radiation moves to a large distance. At the "surface" of a star, it is obvious that $W = \frac{1}{2}$ (actually a little less because of limb-darkening) but in an extended stellar envelope $W \ll 1$ (and in a planetary nebula, $W \sim 10^{-14}$). Thermodynamic equilibrium requires that $W \equiv 1$, so it is clear that detailed balancing in radiative transitions cannot in general occur in a stellar atmosphere.



In addition to being dilute, the stellar radiation field has a markedly non-Planckian frequency distribution. As we know from the Eddington–Barbier relation, the emergent specific intensity at frequency $\nu$ is approximately equal to the source function $S_\nu$ at $\tau_\nu = 1$. Even if $S_\nu$ were $B_\nu$, the fact that the material is vastly more opaque at some frequencies than at others (line to continuum ratios are often $10^3$ and may reach much larger values) implies that the radiation will emerge from greatly differing depths at substantially different temperatures; the radiation field is therefore a composite of widely differing radiation temperatures. The effects of the temperature gradient become extreme when $h\nu/kT > 1$, for then the Planck function varies as $\exp(-h\nu/kT)$ and becomes very sensitive to small changes in $T$. If we were to parameterize the radiation field by introducing a radiation temperature $T_R(\mu, \nu)$ such that for $\mu \geq 0$, $I(r_*, \mu, \nu) = WB_\nu[T_R(\mu, \nu)]$, we would find marked variations of $T_R$ with both $\nu$ and $\mu$. For example, in the solar spectrum, $T_R$ ranges from 4800°K in the visible to ~25,000°K in the ultraviolet in the ground-state continuum of $He^+$. In sum, the radiation field displays an extremely complex behavior, and the conditions required to assure LTE a priori are simply not met.

### THE ELECTRON VELOCITY DISTRIBUTION

In stellar atmospheres, the free electrons are produced by photoionization and collisional ionization. The inverse processes are radiative recombination and three-body collisions, which lead to recaptures of electrons into bound states. While in the continuum, an electron may undergo elastic collisions with other electrons and inelastic collisions (leading to excitation or ionization of bound electrons) with atoms and ions. The elastic collisions redistribute energy among the electrons and tend to lead to an equilibrium partitioning—hence a Maxwellian velocity distribution. If a Maxwellian velocity distribution is in fact attained, we may define the local temperature to be the kinetic temperature of the electrons. On the other hand, inelastic collisions and recombinations disturb the achievement of a Maxwellian velocity distribution, for the inelastic collisions involve electrons only in certain velocity ranges and tend systematically to shift them to much lower velocities, while recombinations remove electrons from the continuum and prevent further elastic collisions. Whether or not the Maxwellian velocity distribution is established hinges upon how rapidly thermalization by elastic collisions occurs compared to the perturbing processes: if it occurs much more rapidly, the velocity distribution will be very nearly Maxwellian.

The thermalization rate can be measured in terms of the *relaxation time* of the system, which, for particles interacting with themselves, is

$$t_c = m^{\frac{1}{2}}(3kT)^{\frac{3}{2}}/[17.9n_e e^4 Z^4 \ln(D/p_0)] \text{ sec} \tag{5-37}$$

(see **598**, Chap. 5). Here $D$ is the *Debye radius* (see §9-4) $D = (kT/8\pi e^2 n_e)^{\frac{1}{2}}$ and $p_0 = e^2/mv^2$ is the impact parameter for a 90° collision. Now consider recombinations; if $\sigma$ is the average cross-section for the process then the mean time between recombinations is

$$t_r = (N\sigma\langle v\rangle)^{-1} = N^{-1}\sigma^{-1}(\pi m/8kT)^{\frac{1}{2}} \text{ sec} \tag{5-38}$$

where $N$ is the density of the particles with which recombination occurs. Two astrophysically important processes are (a) $H + e \rightarrow H^-$ and (b) $H^+ + e \rightarrow H$. At $T \sim 6000°K$ (a typical solar temperature) $\sigma_{H^-} \sim 3 \times 10^{-22}$ cm$^2$ and $n_e/N_H \sim 10^{-4}$. At $T \sim 10000°K$, $\sigma_H \sim 6 \times 10^{-21}$ cm$^2$ and $n_e/n_p \sim 1$. Substituting these values into equations (5-37) and (5-38), we find $t_r/t_c \sim 10^5$ for process (a) and $t_r/t_c \sim 10^7$ for process (b). We conclude, therefore, that under representative conditions in stellar atmospheres a free electron will undergo an *enormous* number of elastic scatterings between recombinations, and that the latter will not seriously hinder equilibration to a Maxwellian distribution.

Let us now consider inelastic collisions. Collisions of electrons with the most abundant element, hydrogen, occur frequently, but the excitation energy of hydrogen is 10 eV while the thermal energy of the electrons is 1 eV. Thus only $3 \times 10^{-5}$ of the electrons have sufficient energy to induce the excitation, and only a fraction of these will be effective. Using typical excitation cross-sections one finds that (at 10,000°K) the rate of inelastic excitations is of the same order as the recombination rate—i.e., very small compared to the elastic collision rate. One must also consider collisions with other elements, which may be grouped as follows: (a) the alkalis, which have large cross-sections but low abundances ($10^{-6}$); (b) Fe, which has numerous low-lying levels and moderate abundance ($4 \times 10^{-5}$); and (c) C, N, and O, which have small cross-sections but large abundance ($10^{-3}$). Most of the levels for groups (b) and (c) are metastable, so that most of the inelastic excitations are subsequently cancelled by collisional de-excitation; ignoring this effect we overestimate the number of inelastic excitations. Taking the various factors into account and ignoring compensating de-excitation, Böhm estimates (elastic collisions/inelastic collisions) $\sim 10^3$ and hence concludes that a Maxwellian velocity distribution that defines $T_e$ is established. Recent work (**573**) suggests that departures from a Maxwellian distribution in a pure hydrogen gas can occur in the high-energy tail if (a) the ionization level is very low ($n_e/n_H \lesssim 0.01$) and (b) the ground-state



population is far from its equilibrium value; these conditions can occur in the solar chromosphere.

Finally, one may ask if the atoms and ions in the atmosphere also have a Maxwellian velocity distribution, and if their kinetic temperature $T_k = T_e$. An analysis of this question (**88**) for a pure hydrogen atmosphere of atoms, ions, electrons, and radiation, demanding a steady-state solution, while allowing for energy exchange among the four components of the medium, shows that if $n_e > 10^{10}$ (a condition easily met in the bulk of the atmosphere) and $5 \times 10^3 < T_e < 10^5$, then $|T_k - T_e| \lesssim 10^{-3} T_e$. It thus appears safe to conclude that a unique local kinetic temperature applies to all the particles in most atmospheric regions.

### THE IONIZATION EQUILIBRIUM

The degree of ionization of stellar material is determined by the balance of photoionizations and collisional ionizations against radiative recombinations and three-body collisional recombinations. Let us first examine the relative rates of photoionization and collisional ionization; it suffices to obtain only an order-of-magnitude estimate.

The energy absorbed by an atom in bound state $i$ at frequency $\nu$ in interval $d\nu$ is $4\pi J_\nu \alpha_i(\nu)\, d\nu$; each photon has energy $h\nu$, hence the total number of photoionizations is

$$n_i R_{ik} = n_i 4\pi h^{-1} \int_{\nu_0}^{\infty} \alpha_i(\nu) J_\nu \nu^{-1}\, d\nu \tag{5-39}$$

To estimate $R_{ik}$, we adopt a hydrogenic cross-section

$$\alpha_\nu = (\pi e^2/mc) f_c (2\nu_0^2/\nu^3)$$

where $f_c$ is the integrated oscillator strength for the continuum. Further, we write

$$J_\nu = W B_\nu(T_R) = W(2h\nu^3/c^2) \sum_{n=1}^{\infty} \exp(-nh\nu/kT_R)$$

Then
$$R_{ik} = (16\pi^2 e^2 \nu_0^2/mc^3) f_c W \sum_{n=1}^{\infty} E_1(nh\nu_0/kT_R) \tag{5-40}$$

The rate of collisional ionizations can be computed from $\sigma(v)$, the collisional ionization cross-section for electrons of velocity $v$:

$$n_i C_{ik} = n_i n_e \int_{v_0}^{\infty} \sigma(v) f(v) v\, dv \tag{5-41}$$

TABLE 5-1
*Ratio of Radiative to Collisional Ionization Rates*

| Star | $\chi_{ion} = 8$eV | $\chi_{ion} = 1$eV |
|---|---|---|
| Sun | $10^3$ | 2 |
| O-star | 20 | 0.2 |

SOURCE: From data by K. H. Böhm, in *Stellar Atmospheres*, ed. J. L. Greenstein, Chicago: University of Chicago Press, 1960, by permission.

To obtain an estimate, we adopt the semiclassical Thomson formula [cf. (**684**, 120)]

$$\sigma(v) = 3f_c\pi e^4E^{-1}[(h\nu_0)^{-1} - E^{-1}] \tag{5-42}$$

where $E = \frac{1}{2}mv^2$ is the energy of the incident electron. Substituting equations (5-2) and (5-42) into (5-41) and integrating we obtain

$$C_{i\kappa} = n_e[12\pi^{\frac{1}{2}}e^4f_c/(2mk^3T_e^{\,3})^{\frac{1}{2}}]u_0^{\,-1}E_2(u_0) \tag{5-43}$$

where $u_0 \equiv h\nu_0/kT_e$. In the limit that $h\nu_0 \gg kT_R$ and $h\nu_0 \gg kT_e$ we retain only the first term of equation (5-40), and use the asymptotic result that for $x \gg 1$, $E_2(x) \to E_1(x) \to e^{-x}/x$ to obtain

$$\frac{R_{i\kappa}}{C_{i\kappa}} \approx \frac{4(2\pi^3k)^{\frac{1}{2}}h\nu_0^{\,3}}{3m^{\frac{1}{2}}e^2c^3}\left(\frac{WT_R}{n_eT_e^{\frac{1}{2}}}\right)\exp\left[h\nu_0\left(\frac{1}{kT_e} - \frac{1}{kT_R}\right)\right] \tag{5-44}$$

For photospheric layers we could adopt $W \approx \frac{1}{2}$, $T_R \approx T_e$. Böhm calculates estimates of $R_{i\kappa}/C_{i\kappa}$ for representative cases of levels with ionization potentials of 1 eV and 8 eV for conditions characteristic of the outer layers ($\tau \sim 0.05$) of the sun and an O-star. In particular, for the sun he adopts $n_e \approx 3 \times 10^{12}$, and $T \approx 5 \times 10^3$ °K while for the O-star he uses $n_e \approx 3 \times 10^{14}$, $T \approx 3.2 \times 10^4$ °K and finds the values for $R_{i\kappa}/C_{i\kappa}$ listed in Table 5-1. It is clear that *in stellar photospheres, the radiative rates dominate*, except for high-lying levels at high temperature and densities. In fact, for O-stars the important levels have even larger values of $\chi_{ion}$ than those listed in Table 5-1 (e.g., the ground state of H at 13.6 eV and the ground state of He I at 24.5 eV), and are even more markedly radiatively dominated. Thus *the ionization equilibrium is vulnerable to departures from LTE if $J_\nu$ departs from $B_\nu$*. Note in passing that in the *corona* of a star where $T_e \sim 2 \times 10^6$ °K and $T_R \sim 6 \times 10^3$ °K (for the sun), while the relevant values of $h\nu_0$ are around 300 eV, the exponential factor in equation (5-44) becomes very small and *collisional ionizations dominate*.



We may carry out similar estimates for the rates of radiative recombination and three-body collisional recombination; these processes are both essentially collisional and hence *per ion* occur at the LTE rate. We can then use detailed balancing arguments to compute the rates in terms of the equilibrium values of the upward rates. We may still use equation (5-44), except that for the radiative recombinations the appropriate temperature is now $T_e$ not $T_R$, and $W \equiv 1$. We then find that radiative recombinations always outweigh collisional recombination, both in the photosphere and corona (in the corona, yet another mechanism—dielectronic recombination—outweighs radiative recombination).

The ionization balance is thus determined by photoionizations and radiative recombination; to establish the equilibrium the numbers of ionizations and recombinations are equal: $n_iR_{i\kappa} = n_\kappa R_{\kappa i} = n_i^*R_{i\kappa}^*$ where the last equality follows by a detailed-balance argument. Hence for the ground state,

$$\begin{aligned}(n_{0j}^*/n_{0j}) &= 4\pi W\int_{\nu_0}^{\infty}(h\nu)^{-1}\alpha_\nu B_\nu(T_R)\,d\nu \Big/ 4\pi\int_{\nu_0}^{\infty}(h\nu)^{-1}\alpha_\nu B_\nu(T_e)\,d\nu \\ &= W\int_{h\nu_0/kT_R}^{\infty}e^{-x}x^{-1}\,dx \Big/ \int_{h\nu_0/kT_e}^{\infty}e^{-x}x^{-1}\,dx \\ &= WE_1(h\nu_0/kT_R)/E_1(h\nu_0/kT_e) \\ &\approx W(T_R/T_e)[\exp(-h\nu_0/kT_R)/\exp(-h\nu_0/kT_e)]\end{aligned} \tag{5-45}$$

where we have again used hydrogenic cross-sections. If we substitute for $n_{0j}^*$ from the Saha equation (5-13) we may obtain the *approximate ionization equation*

$$\begin{aligned}&(n_en_{0,\,j+1}/n_{0,\,j}) \\ &\quad = W\cdot(2g_{0,\,j+1}/g_{0,\,j})(2\pi mkT_R/h^2)^{\frac{3}{2}}\cdot(T_e/T_R)^{\frac{1}{2}}\cdot\exp(-\chi_{Ij}/kT_R)\end{aligned} \tag{5-46}$$

which has been extensively applied—e.g., in analyses of gaseous nebulae.

To analyze the ionization balance in stellar atmospheres we now must decide (a) how to choose $J_\nu$, and (b) which levels dominate. Böhm suggested comparing the values of $4\pi\int(h\nu)^{-1}\kappa_\nu J_\nu\,d\nu$ with $4\pi\int(h\nu)^{-1}\kappa_\nu B_\nu\,d\nu$, where $\kappa_\nu$ is the total opacity from *all* overlapping continua and $J_\nu$ is the mean intensity obtained from LTE model-atmosphere calculations. If these numbers are equal the claim is made that LTE is self consistent. Böhm examines the FeI $\leftrightarrow$ Fe II equilibrium in a model solar atmosphere and finds that the rates mentioned above have a ratio of 2.9 at $\bar{\tau} = 0.01$, 1.3 at $\bar{\tau} = 0.05$, and essentially unity at $\bar{\tau} \geqslant 0.1$; from this one is tempted to conclude that the Saha ionization formula is valid below $\bar{\tau} = 0.1$.

There are, however, flaws in this argument. First, it is clear that integrated rates summed over all continua of an atom may be subject to cancellations

and compensations, and it is not at all clear what a given departure between the two integrals implies for any particular level (i.e., some levels may be overpopulated and others underpopulated and the integrals could balance). Second, and far more important, the reasoning is circular if from the outset we use $B_\nu$ as $S_\nu$ to calculate $J_\nu$, for we know that

$$J_\nu(\tau_\nu) = \Lambda_{\tau_\nu}[B_\nu(t_\nu)] = B_\nu(\tau_\nu) + 0(e^{-\tau_\nu})$$

i.e., $J_\nu$ is *forced* to $B_\nu$ at $\tau_\nu \gtrsim 1$, and the two integrals automatically become equal artificially.

As we shall show below (cf. §7-5 and Chapters 11 and 12) the characteristic feature of non-LTE transfer is that the source function contains a dominant *scattering term* that is only weakly coupled to local thermal parameters. In such cases, $S_\nu$ may differ greatly from $B_\nu$ over *large* ranges of optical depth (essentially to the depth from which a photon cannot escape without being thermalized despite the high probability of scattering and low probability of destruction). Furthermore, if we attempt to find $S_\nu$ by starting with $B_\nu$, computing $J_\nu$, using this to re-evaluate $S_\nu$ and iterating, it is found (cf. §6-1) that the convergence rate is *extremely slow*; with a single iteration one inevitably obtains an estimate of $S_\nu$ that is very close to LTE, *but that is false*. Further iterations (perhaps thousands!) are required to propagate information about the existence of a boundary to the deeper layers by the inefficient iteration procedure; each iteration will show a continuing, progressive departure from LTE, and when *strict* consistency between $S_\nu$ and $J_\nu$ is obtained, the departures are much larger and extend far deeper than the first iteration indicates. In short, experience has shown that estimates of the kind outlined above, based upon a single iteration away from LTE are worthless. The fact that the seemingly plausible arguments based on such estimates are false was not realized in much of the classical work on stellar atmospheres, and erroneous conclusions about the validity of LTE were drawn; we shall return to this crucially important point in Chapters 7, 11, and 12—further discussion has also been given by Thomas (**626**, 141–147).

In summary, we have shown that in the ionization process radiative rates dominate collisional, and given the nonequilibrium character of the radiation field we must expect that LTE will *not* be valid, and therefore from the outset we must perform a simultaneous solution of both the statistical equilibrium and transfer equations. Only when a strictly self-consistent solution is obtained is it possible to decide in which regions LTE actually prevails.

#### THE EXCITATION EQUILIBRIUM

As in the case of ionizations, we again ask whether any particular transition is dominated by collisional or radiative processes. The radiative excitation rate is given by $B_{ij} \int \phi_\nu J_\nu \, d\nu$, which we shall estimate by replacing



TABLE 5-2
*Ratio of Collisional to Radiative Excitation Rates*

| Star | $\lambda$(Å) = 3000 | 4000 | 5000 | 6000 | 7000 | 8000 | 9000 |
|---|---|---|---|---|---|---|---|
| Sun | 0.003 | 0.007 | 0.017 | 0.035 | 0.061 | 0.099 | 0.15 |
| O-Star | 0.19 | 0.44 | 0.85 | 1.4 | 2.2 | 3.1 | 4.2 |

SOURCE: K. H. Böhm, in *Stellar Atmospheres*, ed. J. L. Greenstein, Chicago: University of Chicago Press, 1960, by permission.

$J_\nu$ with $WB_\nu$. We again use equation (5-43) to calculate the collision rate with $f_c$ replaced by $f_{ij} = B_{ij}h\nu mc/4\pi^2e^2$. We then find

$$\frac{C_{ij}}{R_{ij}} = \left[\frac{3e^2m^{\frac{1}{2}}\lambda^3}{2h(2\pi^3kT)^{\frac{1}{2}}}\right]\left(\frac{n_e}{W}\right)E_2(x)(e^x - 1) \tag{5-47}$$

where $x \equiv h\nu_{ij}/kT$. Böhm estimates this ratio for typical conditions in the solar atmosphere and the atmosphere of an O-star (using $W \equiv 1$) and obtains the results shown in Table 5-2. We see that the radiative rates dominate except in the red and infrared of hot stars. The same remarks made above about the non-Planckian nature of the radiation field apply (even more strongly!) in the lines; hence again we conclude that the statistical equilibrium and transfer equations must be solved self-consistently. One might be tempted to conclude that LTE must prevail in the long-wavelength line spectrum on the basis of the dominance of collision rates. But as we shall see in §12-4 this is not true, and, in fact, these lines often show the largest effects of departures from LTE!

## 5-4 The Non-LTE Rate Equations

Let us now consider the *equations of statistical equilibrium* (or *rate equations*) by which we calculate the actual occupation numbers of bound and free states of atoms in stellar atmospheres. We shall make the simplifying assumption of *complete redistribution* in the lines (i.e., *the emission and absorption profiles are taken to be identical*); the equations developed on this basis will be used in our discussion of line formation through Chapter 12, and consideration of partial redistribution effects will be deferred until Chapter 13.

#### GENERAL FORM

Consider a volume element in a moving medium. The number density of particles of a given (bound or free) state $i$ of chemical species $k$ will change in time according to the *net flux* of particles through the volume and the

*net rate* at which particles are brought from other states, $j$, by radiative and collisional processes; that is,

$$(\partial n_{ik}/\partial t) = -\nabla \cdot (n_{ik}\mathbf{v}) + \sum_{j \neq i} (n_{jk}P_{ji} - n_{ik}P_{ij}) \tag{5-48}$$

Here $P_{ij}$ denotes the total rate from level $i$ to $j$. The first term on the right-hand side can be shown, by use of the divergence theorem, to be the net number of particles streaming into and out of the unit volume. If we sum over *all* states of species $k$ and write $N_k = \sum_i n_{ik}$, then we have a *continuity equation* for this species

$$(\partial N_k/\partial t) + \nabla \cdot (N_k\mathbf{v}) = 0 \tag{5-49}$$

Multiplying equation (5-49) by $m_k$, the mass of species $k$, and summing over all chemical species we obtain the standard hydrodynamical continuity equation

$$(\partial \rho/\partial t) + \nabla \cdot (\rho\mathbf{v}) = 0 \tag{5-50}$$

where $\rho = \sum_k m_k N_k$. For a *steady state*, equation (5-48) simplifies to

$$n_{ik} \sum_{j \neq i} P_{ij} - \sum_{j \neq i} n_{jk}P_{ji} = -\nabla \cdot (n_{ik}\mathbf{v}) \tag{5-51}$$

and for a *static* atmosphere we have (suppressing the subscript $k$)

$$n_i \sum_{j \neq i} P_{ij} - \sum_{j \neq i} n_j P_{ji} = 0 \tag{5-52}$$

As we consider only static media through Chapter 13, we shall deal almost exclusively with equation (5-52). We shall show how to handle moving media in Chapter 14, and will mention some implications of the righthand side of equation (5-51) in Chapter 15. The total rate $P_{ij}$ in general contains both radiative and collisional terms; let us now write these out in detail.

### RADIATIVE RATES

(a) *Bound–Bound Transitions.* We shall develop two notations (which have *identical* physical content) by writing bound–bound rates in terms of Einstein transition probabilities or in terms of energy-absorption cross-sections; the former is useful for analytical manipulation with simplified atomic models, while the latter allows bound–bound and bound–free rates to be written in an identical form well-adapted to model-atmosphere computation. Consider transitions from bound level $i$ to a higher bound level $j$, in a line with absorption (and emission) profile $\phi_\nu$. The number of transitions produced by incident intensity $I_\nu$ in the frequency interval $d\nu$ and solid angle $d\omega$ is $n_i B_{ij}\phi_\nu I_\nu \, d\nu \, d\omega/4\pi$ or $n_i(\alpha_{ij}/h\nu)\phi_\nu I_\nu \, d\omega \, d\nu$. In a *static* medium, $\phi_\nu$ is



isotropic, so integrating over all angles and frequencies we have the *number of absorptions* in the line

$$n_i R_{ij} = n_i B_{ij} \int \phi_\nu J_\nu \, d\nu \equiv n_i B_{ij} \bar{J}_{ij} = n_i 4\pi \alpha_{ij} \bar{J}_{ij}/h\nu_{ij}$$
$$= n_i 4\pi \int \alpha_{ij}(\nu)(h\nu)^{-1} J_\nu \, d\nu \tag{5-53}$$

In *moving* media (Chapter 14) we may consider either the *comoving frame*, in which case equation (5-53) remains valid if $J_\nu$ is the mean intensity as measured by an observer at rest with respect to the moving fluid, or the *observer's frame*, in which case $\phi_\nu$ now has an angular dependence, and a double integration over angles and frequencies must be carried out explicitly. Similar remarks apply to other radiative rates given below.

The number of *stimulated emissions* is

$$n_j B_{ji} \int \phi_\nu J_\nu \, d\nu = n_j B_{ji} \bar{J}_{ij} = n_j (g_i B_{ij}/g_j)\bar{J}_{ij} = n_j (4\pi/h\nu_{ij})(g_i \alpha_{ij}/g_j)\bar{J}_{ij} \tag{5-54}$$

The *number of spontaneous emissions* is

$$n_j A_{ji} \int \phi_\nu \, d\nu = n_j (2h\nu_{ij}{}^3/c^2) B_{ji} = n_j (2h\nu_{ij}{}^3/c^2)(4\pi/h\nu_{ij})(g_i \alpha_{ij}/g_j) \tag{5-55}$$

The *total downward rate* is the sum of the spontaneous and stimulated rates:

$$n_j R'_{ji} = n_j (A_{ji} + B_{ji}\bar{J}_{ij}) = n_j (4\pi/h\nu_{ij})(g_i \alpha_{ij}/g_j)[(2h\nu_{ij}{}^3/c^2) + \bar{J}_{ij}] \tag{5-56}$$

A prime has been added to $R'_{ji}$ so as to reserve the unadorned symbol for a different use below. We may rewrite equation (5-56) by factoring out the term $(n_i/n_j)^* = g_i \exp(h\nu_{ij}/kT)/g_j$ from the righthand side, and express the total downward rate as

$$n_j R'_{ji} \equiv n_j \left(\frac{n_i}{n_j}\right)^* R_{ji} = n_j \left(\frac{n_i}{n_j}\right)^* \left[4\pi \int \frac{\alpha_{ij}(\nu)}{h\nu} \left(\frac{2h\nu^3}{c^2} + J_\nu\right) e^{-h\nu/kT} \, d\nu\right] \tag{5-57}$$

At first sight equation (5-57) appears very cumbersome, for clearly the terms $(h\nu)^{-1}$, $\nu^3$, and $\exp(-h\nu/kT)$ can all be taken outside the integral owing to the swift decrease of $\phi_\nu$ away from $\nu_{ij}$. We have written the downward rate in this particular way because it is then of exactly the same form as the downward rate in the continuum; moreover, the downward collision rates will also have a factor of $(n_i/n_j)^*$ appearing explicitly. In the end we achieve notational economy in the full rate equations by using equation (5-57) rather than the simpler Einstein probability form.

Finally, it is sometimes useful to work with the *net rate* from level $j$ to level $i$,

$$n_j(A_{ji} + B_{ji}\bar{J}_{ij}) - n_i B_{ij}\bar{J}_{ij} \equiv n_j A_{ji} Z_{ji} \tag{5-58}$$

where the term $Z_{ji}$ is called the *net radiative bracket* (NRB). Net radiative brackets are useful notational devices that we shall employ in Chapters 11 and 12. Further, $Z_{ji}$ can be rewritten as

$$Z_{ji} = 1 - \bar{J}_{ij}(n_i B_{ij} - n_j B_{ji})/(n_j A_{ji}) \equiv 1 - (\bar{J}_{ij}/S_{ij}) \tag{5-59}$$

where $S_{ij}$ is the frequency-independent line source function. Because the NRB contains only the *ratio* of $\bar{J}$ to $S$, it is often true that it is known to much higher accuracy in an iterative procedure than either $S$ or $\bar{J}$ themselves. Under favorable conditions, use of NRB's can significantly enhance the convergence of certain types of solutions of multilevel line-formation problems. If a particular line $i \rightarrow j$ is in *radiative detailed balance*, then $Z_{ji} \equiv 0$, and we may *cancel* the corresponding terms out of the rate equations *analytically* (i.e., *omit* $R_{ij}$ and $R_{ji}$); this situation occurs when a particular line *thermalizes*, and the cancellation procedure is of great use in simplifying the rate equations (cf. §7-5).

(b) *Bound–Free Transitions.* Let us now calculate the radiative rates from a bound level $i$ to the continuum $\kappa$. Let $\alpha_{i\kappa}(\nu)$ be the photoionization cross-section at frequency $\nu$; then the number of photoionizations is calculated by dividing the energy absorbed in interval $d\nu$ by the appropriate photon energy $h\nu$, and summing over all frequencies. Thus the *number of photoionizations* is

$$n_i R_{i\kappa} = n_i 4\pi \int_{\nu_0}^{\infty} \alpha_{i\kappa}(\nu)(h\nu)^{-1} J_\nu \, d\nu \tag{5-60}$$

We may calculate the *number of spontaneous recombinations* by use of a detailed-balancing argument. In *thermodynamic equilibrium*, the number of spontaneous recombinations must equal the number of photoionizations calculated from equation (5-60) when (a) $J_\nu$ has its equilibrium value (i.e., $B_\nu$) and (b) we correct for stimulated emissions at the T.E. value by multiplying by a factor of $(1 - e^{-h\nu/kT})$ (c.f. §4-3). Thus if $n_\kappa$ denotes the ion density,

$$(n_\kappa R'_{\kappa i})^*_{\text{spon}} = n_i^* 4\pi \int_{\nu_0}^{\infty} \alpha_{i\kappa}(\nu)(h\nu)^{-1} B_\nu (1 - e^{-h\nu/kT}) \, d\nu \tag{5-61}$$

The recombination process is a collisional process involving electrons and ions, and therefore is proportional to $n_\kappa \cdot n_e$. For a given electron density and a given $T_e$, which by definition describes the electron velocity distribution, the rate just calculated above must still apply *per ion*, even out of T.E. Hence to obtain the non-LTE spontaneous recombination rate we need correct equation (5-61) only by using the actual ion density $n_\kappa$. Then

$$\begin{aligned}(n_\kappa R'_{\kappa i})_{\text{spon}} &= n_\kappa (n_i/n_\kappa)^* \cdot 4\pi \int_{\nu_0}^{\infty} \alpha_{i\kappa}(\nu)(h\nu)^{-1} B_\nu (1 - e^{-h\nu/kT}) \, d\nu \\ &= n_\kappa (n_i/n_\kappa)^* 4\pi \int_{\nu_0}^{\infty} \alpha_{i\kappa}(\nu)(h\nu)^{-1} (2h\nu^3/c^2) e^{-h\nu/kT} \, d\nu \end{aligned} \tag{5-62}$$



Recall from equation (5-14) that $(n_i/n_\kappa)^* = n_e \Phi_{i\kappa}(T)$, which shows that the spontaneous recombination rate depends on the product of the electron and ion densities and a function of the temperature (which itself depends on atomic properties through the cross-section).

The number of stimulated recombinations may be calculated by a similar procedure; in T.E.,

$$(n_\kappa R'_{\kappa i})^*_{\text{stim}} = n_i^* \cdot 4\pi \int_{\nu_0}^{\infty} \alpha_{i\kappa}(\nu)(h\nu)^{-1} B_\nu e^{-h\nu/kT} \, d\nu \tag{5-63}$$

To generalize this result for the non-LTE case, we (a) replace the equilibrium radiation field $B_\nu$ by the actual value $J_\nu$, and (b) use the actual ion density $n_\kappa$:

$$(n_\kappa R'_{\kappa i})_{\text{stim}} = n_\kappa (n_i/n_\kappa)^* 4\pi \int_{\nu_0}^{\infty} \alpha_{i\kappa}(\nu)(h\nu)^{-1} J_\nu e^{-h\nu/kT} \, d\nu \tag{5-64}$$

The total number of recombinations is, therefore,

$$\begin{aligned} n_\kappa (R'_{\kappa i,\,\text{spon}} + R'_{\kappa i,\,\text{stim}}) &\equiv n_\kappa (n_i/n_\kappa)^* R_{\kappa i} \\ &= n_\kappa \left(\frac{n_i}{n_\kappa}\right)^* 4\pi \int_{\nu_0}^{\infty} \frac{\alpha_{i\kappa}(\nu)}{h\nu} \left(\frac{2h\nu^3}{c_2} + J_\nu\right) e^{-h\nu/kT} \, d\nu \end{aligned} \tag{5-65}$$

The number of recombinations is sometimes expressed in terms of a *recombination coefficient* $\alpha_{\text{RR}}(T)$ defined such that total recombination rate as given by equation (5-65) is $n_\kappa n_e \alpha_{\text{RR}}(T)$.

By comparison of equations (5-53) and (5-60) we see that in writing complete rate equations we can systematize our notation and write *all upward radiative rates* $(i \rightarrow j)$, *for j bound or free, as* $n_i R_{ij}$ *where*

$$R_{ij} \equiv 4\pi \int_{\nu_0}^{\infty} \alpha_{ij}(\nu)(h\nu)^{-1} J_\nu \, d\nu \tag{5-66}$$

and, by comparison of equations (5-57) and (5-65), *all downward radiative rates* $(j \rightarrow i)$ *as* $n_j (n_i/n_j)^* R_{ji}$ *where*

$$R_{ji} \equiv 4\pi \int_{\nu_0}^{\infty} \alpha_{ij}(\nu)(h\nu)^{-1} [(2h\nu^3/c^2) + J_\nu] e^{-h\nu/kT} \, d\nu \tag{5-67}$$

Note that in equilibrium, $R_{ij}^* = R_{ji}^*$.

### COLLISIONAL RATES

The gas in stellar atmospheres is a plasma consisting of atoms, ions, and electrons, among which a wide variety of collisions may occur and produce excitations and ionizations. In cool stars, where the material is primarily

neutral, collisions with neutral hydrogen atoms are numerous and important. As the material becomes appreciably ionized, however, collisions with charged particles predominate owing to the long-range nature of Coulomb interactions. Moreover, because the collision frequency is proportional to the flux of impinging particles, and hence to their velocity, we normally need to consider only collisions with electrons, for in thermal equilibrium their velocities are a factor of $(m_H A/m_e)^{\frac{1}{2}} \approx 43A^{\frac{1}{2}}$ larger than those of ions of atomic weight $A$.

If we denote the cross-section for producing the transition $(i \to j)$ (where $j$ may be bound or free) by collisions with electrons of velocity $v$ as $\sigma_{ij}(v)$, then the total number of transitions is

$$n_i C_{ij} = n_i n_e \int_{v_0}^{\infty} \sigma_{ij}(v) f(v) v \, dv \equiv n_i n_e q_{ij}(T) \tag{5-68}$$

where $v_0$ is the velocity corresponding to $E_0$, the threshold energy of the process—i.e., $\frac{1}{2}m_e v_0^2 = E_0$. The downward rate $(j \to i)$ can be obtained immediately on the basis of detailed-balancing arguments, for the electron velocity distribution is the equilibrium (i.e., Maxwellian) function; thus we must have

$$n_i^* C_{ij} \equiv n_j^* C_{ji} \tag{5-69}$$

from which it follows that the number of downward transitions is

$$n_j C_{ji} = n_j (n_i/n_j)^* C_{ij} = n_j (n_i/n_j)^* n_e q_{ij}(T) \tag{5-70}$$

As was the case for radiative transitions, it is sometimes useful to introduce the *net collisional bracket* $Y_{ij}$ and write the net rate for collisions $i \to j (E_i < E_j)$ as

$$n_i C_{ij} Y_{ij} \equiv n_i C_{ij} - n_j C_{ji} = n_i C_{ij} [1 - (n_j/n_j^*)(n_i^*/n_i)] \tag{5-71}$$

The actual cross-sections $\sigma_{ij}$ required to compute rates are found either experimentally or by rather complicated quantum-mechanical calculations; it would take us too far afield to describe these methods here, so we merely refer the interested reader to (**410**). There exists a vast literature containing results (theoretical and experimental) for a variety of transitions of astrophysical interest; bibliographies of this literature are issued from time to time by the Information Center of the Joint Institute for Laboratory Astrophysics of the University of Colorado and the National Bureau of Standards. (This center also maintains current literature references in an on-line computer.) As indicated by equation (5-68) we are more directly interested in *rates* for a given cross-section, so let us examine $q_{ij}$ in a bit more detail. Usually cross-sections are measured in units of $\pi a_0^2$, where $a_0$ is the Bohr radius; i.e., we write $\sigma_{ij} = \pi a_0^2 Q_{ij}$. Also, $Q_{ij}$ is usually tabulated in terms of the energy of the exciting particle, so writing $\frac{1}{2}mv^2 = E$, and substituting



equation (5-2) into (5-68), we find

$$q_{ij}(T) = C_0 T^{\frac{1}{2}} \int_{u_0}^{\infty} Q_{ij}(ukT) u e^{-u} \, du \tag{5-72}$$

where $u \equiv E/kT$, and $C_0 = \pi a_0^2 (8k/m\pi)^{\frac{1}{2}} = 5.5 \times 10^{-11}$. Writing $x = (u - u_0)$, where $u_0 = E_0/kT$, we obtain

$$q_{ij}(T) = C_0 T^{\frac{1}{2}} \exp(-E_0/kT) \Gamma_{ij}(T) \tag{5-73}$$

where

$$\Gamma_{ij}(T) \equiv \int_0^{\infty} Q_{ij}(E_0 + xkT)(x + u_0) e^{-x} \, dx \tag{5-74}$$

*Exercise 5-5*: Verify equations (5-72) through (5-74).

The advantage of writing the collision rate as in equation (5-73) is that the principal sensitivity to the temperature has been factored out in the product $T^{\frac{1}{2}} \exp(-E_0/kT)$ while $\Gamma_{ij}(T)$ is a slowly-varying function of $T$.

Of course the main problem in application is to obtain reliable values of $Q_{ij}$. A characteristic difficulty for astrophysical work is that for many transitions of interest, $kT \ll E_0$, so that the rate depends extremely sensitively upon values of $Q_{ij}$ near threshold. Unfortunately, for $E \approx E_0$ a great computational effort is required to obtain accurate cross-sections because the simplifying approximations that are valid for $E \gg E_0$ break down, and because complicated variations of $Q_{ij}$ result from *resonances* in the collision process. When values for $Q_{ij}$ can be obtained, one typically fits them by numerical procedures to simple analytical approximants, against which the integration in equation (5-74) can be performed analytically.

For the astrophysically important spectra of H, He I, and He II, accurate experimental cross-sections exist for excitation and ionization from the ground state. For transitions arising from excited states one must rely upon theoretical calculations. For many atoms and ions of interest there may be no detailed estimates whatever available, and one must have recourse to rough methods to estimate rates. A very useful (though quite approximate) expression for excitation rates in radiatively permitted transitions can be written (**639**) in terms of the oscillator strength $f_{ij}$, namely

$$C_{ij} = C_0 n_e T^{\frac{1}{2}} [14.5 f_{ij} (I_H/E_0)^2] u_0 \exp(-u_0) \Gamma_e(u_0) \tag{5-75}$$

where $u_0 \equiv E_0/kT$, $I_H$ is the ionization energy of hydrogen, and *for ions*

$$\Gamma_e(u_0) \equiv \max[\bar{g}, 0.276 \exp(u_0) E_1(u_0)] \tag{5-76}$$

The parameter $\bar{g}$ is about 0.7 for transitions of the form $nl \to nl'$, and about 0.2 for transitions of the form $nl \to n'l'$, $n' \neq n$ (**95**). For neutral atoms $\Gamma_e(u_0)$ has a different form [see (**47**)]. It is worth stressing that equations

(5-75) and (5-76) provide, at best, *rough* values and should be applied with caution. In particular, collisions are *not* restricted by the dipole transition selection rules $\Delta l = \pm 1$, and cross-sections for other values of $\Delta l$ may be as large as for $\Delta l = \pm 1$ despite $f_{ij}$ being zero in the dipole approximation. For collisional ionizations there exists a semi-empirical formula (**402**)

$$\sigma_{i\kappa}(E) = \pi a_0^2[2.5\zeta(I_H/E_0)^2] \ln(E/E_0)\{1 - b \exp[-c(E - E_0)/E_0]\}/(E/E_0) \tag{5-77}$$

which yields a rate

$$C_{i\kappa} = C_0 n_e T^{\frac{1}{2}}[2.5\zeta(I_H/E_0)^2]u_0[E_1(u_0) - be^c u_0 E_1(u_1)/u_1] \tag{5-78}$$

where $\zeta$, $b$, and $c$ are empirical quantities fitted to individual atoms, and $u_1 \equiv u_0 + c$. An alternative approximate formula can be obtained by expressing the collisional ionization cross-section in terms of the photoionization cross-section (**73**, 374) which yields a rate (**334**, 121)

$$C_{i\kappa} = 1.55 \times 10^{13}\, n_e T^{-\frac{1}{2}} \bar{g}_i \alpha(\nu_0) \exp(-u_0)/u_0 \tag{5-79}$$

where $\alpha(\nu_0)$ is the threshold photoionization cross-section, and $\bar{g}_i$ is of order 0.1, 0.2, and 0.3 for $Z = 1$, 2, and $>2$, respectively (here $Z$ is the charge on the ion). The same caveats expressed about equation (5-75) apply to equations (5-78) and (5-79) as well.

### AUTOIONIZATION AND DIELECTRONIC RECOMBINATION

In complex atoms with several electrons, the ionization potential is determined by the *lowest* energy to which a sequence of bound states with only *one* excited electron converge (to the *ground state* of the ion plus a free electron). If *two* electrons are excited within the atom, they will, in general, give rise to states with energies both below and above the ionization potential defined above. Subject to certain selection rules [(**172**, 371; **297**, 173)] the states above the ionization limit may *autoionize* to the ground state of the ion plus a free electron. The inverse process is also possible and, if an ion in the ground state suffers a collision with an electron of sufficiently great energy, then a doubly excited state of the atom may be formed. In general, this process will be of little interest because the compound state will immediately autoionize again (typical autoionization transition probabilities $A_a$ are in the range $10^{13}$–$10^{14}$!), and its equilibrium population will be small. In some cases, however, a stabilizing transition occurs in which one of the two excited electrons (usually the one in the lower quantum level) decays radiatively to the lowest available quantum state, leaving a bound atom with a single excited electron. This process can provide an efficient recombination mechanism referred to as *dielectronic recombination*.



In particular, for an ion of chemical species X and charge $Z$, we consider processes of the type

$$X^{+(Z)}(n, l) + e(E, l'' + 1) \leftrightharpoons X^{+(Z-1)}(n', l + 1; n'', l'') \tag{5-80a}$$

followed by the stabilizing transition

$$X^{+(Z-1)}(n', l + 1; n'', l'') \rightarrow X^{+(Z-1)}(n, l; n'', l'') + h\nu \tag{5-80b}$$

which leaves the ion $(Z - 1)$ in a bound excited state. As an example, for $He^+$ we might have

$$He^+(1s) + e(E, l'' + 1) \rightarrow He^0(2p; n''l'')$$
$$He^0(2p; n''l'') \rightarrow He^0(1s; n''l'') + h\nu$$

If we denote the doubly excited state by $d$, the final bound state of the ion $(Z - 1)$ by $b$, and the ground state of the ion $Z$ as $\kappa$, then the number of dielectronic recombinations to state $b$ from $d$ can be written as $n_\kappa R_{db} = n_d A_s$ where $A_s$ is the spontaneous transition probability for the stabilizing emission; to a good approximation (particularly for large $n''$) $A_s = A(n', l + 1; n, l)$ *for the Z ion.* In the limit of low radiation fields, the reverse process in equation (5-80b) can be ignored, and if $A_a$ measures the transition probability for *autoionization*, $n_d$ can be written (**73**, 258) in terms of its equilibrium population $n_d^*$

$$n_d = n_d^* A_a/(A_a + A_s) \tag{5-81}$$

where
$$n_d^* = n_\kappa n_e (g_d/g_\kappa) C_I T^{-\frac{3}{2}} \exp(-\chi_{d\kappa}/kT) \equiv n_\kappa n_e \Phi_{\kappa d}(T) \tag{5-82}$$

Here $\chi_{d\kappa}$ is the energy of state $d$ above the ionization limit.

*Exercise 5-6*: Derive equation (5-82) by applying the Saha equation (5-14) between the continuum $\kappa$ and bound state $b$, and the Boltzmann equation (5-5) between states $b$ and $d$.

Thus from *each* state $d$ to *each* state $b$ we have the number of dielectronic recombinations

$$n_\kappa R_{db} = n_\kappa n_e \Phi_{\kappa d}(T) A_s A_a/(A_s + A_a) \tag{5-83}$$

As in the case of radiative recombinations, it is often useful to define a dielectronic recombination coefficient $\alpha_{DR}$ such that $n_\kappa R_{db} = n_\kappa n_e \alpha_{DR}$. Note that the ratio of numbers of dielectronic to radiative recombinations depends only on the ratio $\alpha_{DR}/\alpha_{RR}$, and hence is independent of density and is a function of temperature only.

Dielectronic recombination plays an important role in stellar atmospheres in two contexts. First, we can calculate a *total* rate for dielectronic recombination by summing over *all* states $(n, l)$ and $(n'', l'')$; as was shown in a classic paper by Burgess (**118**), this process becomes extremely important at high temperatures. For example, at temperatures $T \gtrsim 10^6$ °K, the dielectronic recombination rate for $He^+$ exceeds the radiative recombination rate by a factor of $10^2$ but drops below radiative recombination for $T \lesssim 10^5$ °K. Burgess convincingly demonstrated that dielectronic recombination is the dominant recombination process in the solar corona (where $T \approx 2 \times 10^6$ °K and $n_e \sim 10^8$), and that *this mechanism establishes the coronal ionization balance*. In calculations for the total rate one must sum over vast numbers of states, and the most important contributions come from states with $n'' \gg n'$ and $l'' \gg l + 1$. It is because these high states have large values of $\chi_{d\kappa}$ that high temperatures are required to overcome the exponential factor in equation (5-82) and produce significant dielectronic recombination [i.e., a large value of $kT$ is required for the electrons to have sufficient energy for the reaction in equation (5-80a) to occur]. In the sum one encounters a divergence problem at large $n''$, similar to that found for partition functions, unless one includes *both* probabilities $A_s$ and $A_a$ in equation (5-83) and takes into account the fact that $A_s$ will dominate $A_a$ for sufficiently large $n''$. The calculations of total rates require the estimation of large numbers of stabilization transition probabilities and collision cross-sections (to calculate $A_a$ by detailed balancing arguments), and are at best difficult; an approximate general formula that provides tolerably accurate values for most ions of interest has been developed (**119**).

A further study (**121**) has shown that in the corona the effects of inverse transitions in equation (5-80b) produced by incident photospheric radiation are small, and that the doubly excited states can be destroyed by collisional ionization if densities are sufficiently high, from which one concludes that dielectronic recombination does not play an important role in the deeper layers of the atmosphere (e.g., in the solar chromosphere and photosphere).

A second situation where dielectronic recombinations are important arises for some ions (e.g., C III and N III) that have *low-lying* double-excitation states ($\chi_{d\kappa} < kT$) that feed free electrons into selected bound states. A striking example is afforded by the $2s2p(^1P^0)\ 3d$ state of N III, which lies only 1.6 eV above the ionization potential to N IV $2s^2\ ^1S$, and which feeds electrons directly into the N III $2s^2\ 3d$ levels and thereby produces the famous N III $\lambda$4634–40 ($3d \rightarrow 3p$) emission lines in Of-stars (**115**; **429**; **440**). In such cases one finds that $A_a \gg A_s$, so that $n_d$ is given by its equilibrium value (relative to *actual* ion densities), equation (5-82); furthermore one need sum over only a few states. On the other hand, these processes occur deep enough in the atmosphere that one must account for the inverse transitions



produced by the radiation field in the stabilizing transition. If state $d$ is characterized as $(n', l', L')$ and state $b$ as $(n, l, L)$ we have the total dielectronic recombination rate

$$n_\kappa R_{db} = n_\kappa n_e \sum_{n', l', L', L} \Phi_{n'l'L'}(T) A(n', l', L'; n, l, L)[1 + (2h\nu^3/c^2)^{-1} \bar{J}_{n'l'L'}] \tag{5-84}$$

where $\bar{J}$ is the mean intensity in the stabilizing transition. Often the double excitation state is so broad (because of the very short lifetime against autoionization) that the radiation field used to compute $\bar{J}$ can be fixed at the continuum value. If the $L'$ dependence of $\chi_{n', l', L'}$ is small, then one can define $A^*(n', l'; n, l) \equiv g_{l'}^{-1} \sum_{L', L} g(L') A(n', l, L'; n, l, L)$ and replace $\Phi_{n'l'L'}$ with $\Phi_{n'l'}$ to obtain

$$n_\kappa R_{db} = n_\kappa n_e \sum_{n', l'} \Phi_{n'l'}(T) A^*(n', l'; n, l)[1 + (2h\nu^3/c^2)^{-1} \bar{J}_{n'l'}] \tag{5-85}$$

while the inverse (upward) rate $b \rightarrow d$ becomes

$$n_b R_{bd} = n_b \sum_{n', l'} B^*(n, l; n', l') \bar{J}_{n'l'} \tag{5-86}$$

where $B^*(n, l; n', l') \equiv A^*(n', l'; n, l)(g_{l'} c^2 / 2h\nu^3 g_l)$.

### COMPLETE RATE EQUATIONS

Having examined all of the processes of interest, we may now assemble the individual rates into a single complete equation of the form of equation (5-52) for each bound state $i$ of each ionization stage of each chemical species in the material. We shall (a) ignore explicit mention of dielectronic recombination because the rate has the same *form* as for radiative recombination, and we shall assume that both are included; and (b) assume that all ionizations from bound states of ion $j$ go to the ground state only of ion $j + 1$ (generalization is easy but complicates the notation and discussion). We then may write

$$-\sum_{i'' < i} n_{i''}(R_{i''i} + C_{i''i}) + n_i \left[ \sum_{i'' < i} (n_{i''}/n_i)^*(R_{ii''} + C_{i''i}) + \sum_{i' > i}^{\kappa} (R_{ii'} + C_{ii'}) \right]$$

$$- \sum_{i' > i}^{\kappa} n_{i'}(n_i/n_{i'})^*(R_{i'i} + C_{ii'}) = 0 \tag{5-87}$$

where the radiative rates are defined by equations (5-66) and (5-67), the collision rates by equation (5-68), and the LTE population ratios by equations (5-5) and (5-14) for bound states, and bound and free states, respectively. One such equation may be written for each bound state. We have one more variable (the ion density $n_\kappa$) than we have equations. If we wrote down an ionization equation

$$\sum_{i<\kappa} n_i(R_{i\kappa} + C_{i\kappa}) - n_\kappa \sum_{i<\kappa} (n_i/n_\kappa)^*(R_{\kappa i} + C_{i\kappa}) = 0 \qquad (5\text{-}88)$$

we would find it to be *redundant* with the set (5-87).

*Exercise 5-7*: Show that equation (5-88) results from summing equation (5-87) over all bound states.

We therefore invoke an additional physical constraint to complete the system. For an impurity species (i.e., $\alpha_k/\alpha_H \ll 1$), we close the system by demanding that the total number of atoms and ions (of all kinds) of the species equal the correct fraction of the number of all hydrogen atoms (including protons); i.e.,

$$\sum_{i,j} n_{ij,k} - (\alpha_k/\alpha_H)\left(\sum_i n_{i,H} + n_p\right) = 0 \qquad (5\text{-}89)$$

Alternatively, we can close the system by invoking charge conservation (saving the total number conservation for use elsewhere) and write

$$\sum_k \sum_j jN_{jk} + n_p = n_e \qquad (5\text{-}90)$$

where $N_{jk} = \sum_i n_{ijk}$. The final system, for all levels of all ions of all species is written in the general form

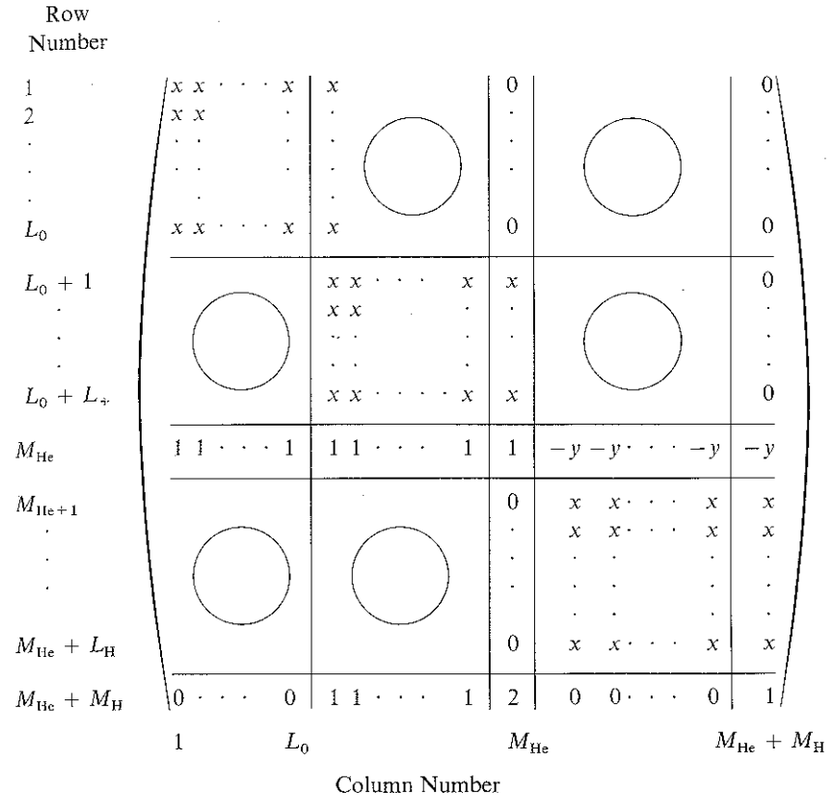
$$\mathscr{A}\mathbf{n} = \mathscr{B} \qquad (5\text{-}91)$$

where $\mathbf{n}$ denotes a vector that lists *all* occupation numbers (say $\mathscr{N}$ of them) while $\mathscr{A}$ is an $(\mathscr{N} \times \mathscr{N})$ matrix and $\mathscr{B}$ is a vector in which only one element is nonzero [from equation (5-89) or (5-90)].

To make these considerations more definite, let us consider a case that is simple enough to be manageable and complicated enough to be of general applicability. Suppose we have an atmosphere composed entirely of hydrogen and helium (of abundance $Y$, by number, relative to hydrogen). We consider the helium to consist of a ladder of three ionization stages, $He^0$, $He^+$, and $He^{++}$, and we suppose that these ions have $L_0$, $L_+$, and 1 levels respectively. Further we write $M_{He} = L_0 + L_+ + 1$; i.e., $M_{He}$ is the total number of helium states of all kinds. Similarly we consider $L_H$ bound states



of hydrogen, yielding $M_H = L_H + 1$ states in all (including protons). Then, using $X$'s to denote nonzero elements, the rate matrix $\mathscr{A}$ has the form




The first $L_0$ rows correspond to equation (5-87) for $He^0$, the next $L_+$ rows give equation (5-87) for $He^+$, the $M_{He}$th row gives the abundance equation (5-89), the next $L_H$ rows give equation (5-87) for H, and the last row gives charge conservation. The vector $\mathbf{n}$ consists of elements

$$\mathbf{n} = [n_1(He^0), \ldots, n_{L_0}(He^0), n_1(He^+), \ldots, n_{L_+}(He^+), n(He^{++}), n_1(H), \ldots, n_{L_H}(H), n_p]^T \qquad (5\text{-}92)$$

and

$$\mathscr{B} = (0, \ldots, 0, n_e)^T \qquad (5\text{-}92)$$

For *given* values of $n_e$, $T$, and the radiation field, equation (5-91) is a *linear system* in $\mathbf{n}$, and may be solved by standard numerical methods (**526**, Chapter 9).

## 5-5 The Non-LTE Equation of State

From the results of the preceding sections, we see that in LTE each occupation number at a specific point in the atmosphere is a function of only *two* thermodynamic variables; i.e., $n_i = n_i(N, T)$ where $T$ is the *absolute* temperature at that point. In contrast, in the non-LTE case, the full rate equations imply that $n_i = n_i(N, T, J_\nu)$ where $J_\nu$ denotes the frequency dependence of the radiation field over the entire spectrum and $T$ is now a *kinetic* temperature describing only the particle velocity distribution function. We now have as many new (fundamental!) thermodynamic variables as are required to specify the *distribution* of radiation in frequency. [Note that *if* we could simplify the description of this distribution—e.g., *if* we could write $J_\nu \equiv WB_\nu(T)$—then the situation would be vastly simplified; but in general we may need to consider perhaps hundreds of new variables.] As was the case for the LTE equation of state, the non-LTE statistical equilibrium equations are actually nonlinear in the electron density $n_e$, and we shall require a linearization procedure to solve for the occupation numbers; but now we shall have to extend the linearization to include changes in the radiation field as well. We shall see in §7-5 that this approach provides a method for coupling the transfer equations and statistical equilibrium equations together, and allows us to determine the global response of the gas to the radiation field simultaneously with the reciprocal response of the radiation field to material properties.

Before developing the linearization procedure required in the general case, it is worthwhile to consider a few examples that illustrate clearly the essential physical content of the statistical equilibrium equations.

### LIMITING CASES

Consider first an atom consisting of a single bound level that can ionize to its continuum. We then have one rate equation which states that (ignoring stimulated emissions)

$$(n_1/n_1^*) = \left[4\pi \int_{\nu_0}^{\infty} (\alpha_\nu B_\nu/h\nu)\, d\nu + n_e q_{1\kappa}\right] \Big/ \left[4\pi \int_{\nu_0}^{\infty} (\alpha_\nu J_\nu/h\nu)\, d\nu + n_e q_{1\kappa}\right] \tag{5-94}$$

We note first that, as the electron density becomes very large, so that collisional rates exceed the radiative rates, then

$$\lim_{n_e \to \infty} (n_1/n_1^*) = \lim_{n_e \to \infty} (n_e q_{1\kappa}/n_e q_{1\kappa}) \equiv 1$$

i.e., LTE is recovered. Further, at very large optical depth, $J_\nu \to B_\nu$ and clearly $n_1/n_1^* \to 1$; i.e., *if* the radiation field is perfectly Planckian we recover



LTE, as expected. Two comments are necessary here, however. (a) To obtain LTE in a *multilevel* atom $J_\nu$ must equal $B_\nu$ in *all* transitions. If *any* transition is transparent, then LTE will *not* be obtained (unless densities are so high that collisions dominate), not only for the particular levels involved in the transition under consideration, but actually for *all* other levels as well because the radiation field in each transition influences the populations of *all* levels (see below). (b) We have left unanswered the question of how large is "very large" optical depth. As we have indicated earlier, $\tau_\nu \gtrsim 1$ is *not* sufficient to guarantee $J_\nu \to B_\nu$. Rather $\tau_\nu$ must exceed a *thermalization depth*, for which precise estimates will be given in Chapters 7 and 11. In the *low-density* limit (e.g., in a nebula), equation (5-94) reduces to

$$(n_1/n_1^*) = \int_{\nu_0}^{\infty} (\alpha_\nu B_\nu/h\nu)\, d\nu \Big/ \int_{\nu_0}^{\infty} (\alpha_\nu J_\nu/h\nu)\, d\nu \tag{5-95}$$

which states that, if the recombination rate exceeds the photoionization rate, the level is overpopulated; and it is underpopulated if the reverse is true. Equation (5-95) is, of course, equivalent to equation (5-46) which is often applied in nebular analyses. In the *coronal* case, we have $T_e(\sim 10^6\ ^\circ\mathrm{K}) \gg T_R(\sim 6000^\circ\mathrm{K})$, which implies that collisional ionizations exceed radiative [see equation (5-44) and related discussion] while radiative plus dielectronic recombinations, both of which proceed at a rate specified by $T_e$, exceed collisional recombinations. Then

$$n_1 n_e q_{1\kappa} = n_\kappa n_e(\alpha_{\mathrm{RR}} + \alpha_{\mathrm{DR}})$$

so that

$$(n_\kappa/n_1) = q_{1\kappa}/(\alpha_{\mathrm{RR}} + \alpha_{\mathrm{DR}}) = f(T) \tag{5-96}$$

That is, the coronal ionization balance depends only on temperature and is independent of the electron density, a fact that vastly simplifies analysis of the corona. Both the coronal and nebular situations represent extreme departures from LTE.

Let us now consider some multilevel problems. Suppose we have a volume of pure hydrogen gas illuminated by a very dilute radiation field (i.e., a nebula). We anticipate that virtually all of the hydrogen will be in its ground state, and we assume that all the resonance lines are completely opaque (and hence in detailed balance). Further, we assume that, after an atom is photoionized from the ground state, recombinations occur to *all* states, but the populations of the upper states are so small and the incident radiation field so diluted that (a) we can ignore photoionization out of these states, and (b) electrons in any excited state cascade downward at rates determined by the Einstein coefficients $A_{ji}$ without reabsorption upward (i.e., the subordinate lines are *transparent*). We further assume densities are so low that we

may neglect collisions. Then we have an ionization equation

$$n_1 R_{1\kappa} = n_e^2 \sum_{i=1}^{I} \alpha_{\mathrm{RR}}(i, T) \tag{5-97a}$$

and a number conservation equation

$$\sum_{i=1}^{I} n_i + n_e = n_{\mathrm{H}} \tag{5-97b}$$

where $n_{\mathrm{H}}$ is the (given) hydrogen density, $I$ is the total number of bound states considered, and $n_e = n_p$ (for pure hydrogen). $R_{1\kappa}$ is assumed given in terms of $J_\nu = WB_\nu(T_R)$, as in equation (5-40). For any subordinate state we can calculate the population in terms of the *branching ratios* $a_{ji} \equiv A_{ji}/\sum_{l<j} A_{jl}$, and the *cascade probabilities* $p_{ji}$ which are defined recursively as $p_{i+1,i} \equiv a_{i+1,i}$, and $p_{ji} \equiv a_{ji} + \sum_{k=i+1}^{j-1} p_{jk} a_{ki}$ for $(j = i + 2, \ldots, I)$. Then for level $i$ we find

$$\begin{aligned} n_i \sum_{l<i} A_{il} &= n_e^2 \alpha_{\mathrm{RR}}(i, T) + \sum_{j>i}^{I} A_{ji} n_j \\ &= n_e^2 \left[ \alpha_{\mathrm{RR}}(i, T) + \sum_{j>i}^{I} p_{ji} \alpha_{\mathrm{RR}}(j, T) \right] \end{aligned} \tag{5-98}$$

*Exercise 5-8*: (a) Verify the expressions for $p_{ji}$ given above and derive equation (5-98). (HINT: Start with level I and work downward.) (b) Show that equations (5-97) and (5-98) yield a quadratic equation in $n_e$ that allows the determination of $n_e(n_H, T)$, and hence all the $n_i$'s. (c) Show that $p_{j1} = 1$ $(j > 1)$; interpret this result physically.

From equation (5-98) we may estimate ratios of occupation numbers, and hence ratios of line intensities along a series. For example we can compute the relative intensities of the Balmer lines (the *Balmer decrement*) as

$$I(\mathrm{H}_k)/I(\mathrm{H}_j) = (n_k A_{k2} h\nu_{k2})/(n_j A_{j2} h\nu_{j2})$$

and compare the theoretical results with observation. The approach outlined in equations (5-97) and (5-98) (with extensive elaboration and refinement!) forms the basis for the analysis of nebulae [see (**15**, Chaps. 23–25; **10**, Chap. 4; **415**, pp. 40–110; and **350**, Chaps. 1–3)].

Finally, consider an atom that consists of three states (1, 2, 3) in order of increasing energy in a rarefied medium (neglect collisions) and a dilute radiation field. A famous result regarding such a system is *Rosseland's theorem of cycles*, which states that the number of radiative transitions in the direction $1 \to 3 \to 2 \to 1$ exceeds the number in the inverse direction $1 \to 2 \to 3 \to 1$. A consequence of this result is that energetic photons are systematically



*degraded* from high energies (say far ultraviolet) to low (visible and infrared); for example, in a nebula, Lyman continuum photons are degraded—e.g., into Balmer continuum photons plus L$\alpha$ photons (state 1 = 1*s*, state 2 = 2*p*, state 3 = continuum). We may calculate the ratio $R_{1\to3\to2\to1}/R_{1\to2\to3\to1}$ quite easily. The number of excitations $1 \to 3$ is $n_1 B_{13} WB(\nu_{13})$. Of the excited atoms in state 3, a fraction $A_{32}/(A_{32} + A_{31})$ decays to state 2, and of the atoms in state 2 a fraction $A_{21}/[A_{21} + B_{23} WB(\nu_{23})]$ decays to state 1 (here we have ignored stimulated emission). Thus

$$n_1 R_{1\to3\to2\to1} = \frac{n_1 B_{13} WB(\nu_{13}) A_{32} A_{21}}{(A_{32} + A_{31})[A_{21} + B_{23} WB(\nu_{23})]} \tag{5-99}$$

By similar reasoning

$$n_1 R_{1\to2\to3\to1} = \frac{n_1 B_{12} WB(\nu_{12}) B_{23} WB(\nu_{23}) A_{31}}{[A_{21} + B_{23} WB(\nu_{23})](A_{32} + A_{31})} \tag{5-100}$$

so that

$$\begin{aligned} R_{1\to2\to3\to1}/R_{1\to3\to2\to1} \\ = W[B_{12}B(\nu_{12})/A_{21}][B_{23}B(\nu_{23})/A_{32}][A_{31}/B_{13}B(\nu_{13})] \end{aligned} \tag{5-101}$$

But using the relations among the Einstein coefficients and writing $B_\nu$ in the Wien approximation $(h\nu/kT \gg 1)$ one finds $[B_{ij}B(\nu_{ij})/A_{ji}] = (n_j/n_i)^*$, so equation (5-101) reduces to $R_{1\to2\to3\to1}/R_{1\to3\to2\to1} = W < 1$, which proves the theorem. The result clearly follows from the fact that in the cycle $1 \to 3 \to 2 \to 1$ the dilution factor enters only *once*, while in the reverse process it enters *twice*. In stellar atmospheres, Rosseland's theorem is relevant because at certain depths one may have resonance lines that are opaque (i.e., $W = 1$) exciting atoms to upper states, from which the subordinate lines are transparent; in such cases we anticipate a systematic photon degradation.

### LINEARIZATION

As mentioned before, the general system $\mathscr{A}\mathbf{n} = \mathscr{B}$ can be solved as a linear system for $\mathbf{n}$ *if* $n_e$, $T$, and $J_\nu$ are all specified. But, in practice, we do *not* know exact values for these variables in the course of a model-atmosphere computation (recall the discussion for the LTE equation of state) but have only *current estimates* in an overall *iterative* process. We expect all of these variables to change by amounts $\delta n_e$, $\delta T$, $\delta J_\nu$, etc. to satisfy better the constraints of energy and pressure balance, and must evaluate the response of $\mathbf{n}$ to these changes, in the form

$$\boldsymbol{\delta}\mathbf{n} = (\partial\mathbf{n}/\partial n_e)\,\delta n_e + (\partial\mathbf{n}/\partial T)\,\delta T + \sum_{k=1}^{K} (\partial\mathbf{n}/\partial J_k)\,\delta J_k \tag{5-102}$$

Here $J_k (k = 1, \ldots, K)$ is the mean intensity at discrete frequencies that sample the spectrum. These frequencies are chosen such that all integrals over frequency are replaced by *quadrature sums*—i.e.,

$$\int F(\nu)\, d\nu = \sum_{k=1}^{K} w_k F(\nu_k) \tag{5-103}$$

We obtain equation (5-102) by linearization of the original equations (5-91) [and also can find parallel linearized equations that give, in essence, $\delta J_k (\delta T, \delta n_e, \boldsymbol{\delta n})$ from the transfer equations; cf. §7-5]. Equations of the form (5-102) are required in two contexts: (a) model atmosphere calculations where all variables may change in an iteration cycle, and (b) multilevel statistical equilibrium calculations for a *given* model ($n_e$, $T$, and total particle density *fixed*). The procedure for case (b) will be deferred until Chapter 12, and we shall consider only case (a) here. [In case (b) one may use a special technique motivated by consideration of the computational methods of solving transfer equations, to be developed in Chapter 6.]

If $x$ denotes any variable, then by linearization of equation (5-91) we have

$$\frac{\partial \mathbf{n}}{\partial x} = \mathscr{A}^{-1}\left[\frac{\partial \mathscr{B}}{\partial x} - \left(\frac{\partial \mathscr{A}}{\partial x} \cdot \mathbf{n}\right)\right] \tag{5-104}$$

where we have assumed that $\mathbf{n}$ *is* the solution of the *current system* $\mathscr{A}\mathbf{n} = \mathscr{B}$ (we might introduce a subscript zero, or some similar device, but it would become unwieldy). An extremely important feature of this approach is that every derivative in equations (5-102) and (5-104) can be written down *analytically* (though the inverse $\mathscr{A}^{-1}$ must be computed numerically); this produces a system of high accuracy and reliability. To illustrate the procedure, we shall write down some representative derivatives for the model atoms discussed at the end of §5-4; more comprehensive collections of formulae are given in (**42**) and (**437**). In what follows, we use the auxiliary vector $\mathbf{a} \equiv (\partial \mathscr{A}/\partial x) \cdot \mathbf{n}$. Suppose we choose some frequency $\nu_k$, and wish to calculate $\partial \mathbf{n}/\partial J_k$. Except for the "special" rows of $\mathscr{A}$ that express abundance and charge conservation, we will generally have

$$(\partial \mathscr{A}/\partial J_k)_{ij} = -4\pi w_k \alpha_{ji}(\nu_k)/h\nu_k, \quad (j < i) \tag{5-105a}$$

$$(\partial \mathscr{A}/\partial J_k)_{ii} = (4\pi w_k/h\nu_k)\left[\sum_{j>i} \alpha_{ij}(\nu_k) + \sum_{j<i} \alpha_{ji}(\nu_k)(n_j/n_i)^* e^{-h\nu_k/kT}\right] \tag{5-105b}$$

and

$$(\partial \mathscr{A}/\partial J_k)_{ij} = -[4\pi w_k \alpha_{ij}(\nu_k)/h\nu_k](n_i/n_j)^* e^{-h\nu_k/kT}, \quad (j > i) \tag{5-105c}$$



from which we find

$$\begin{aligned} a_i &= [(\partial \mathscr{A}/\partial J_k) \cdot \mathbf{n}]_i \\ &= (4\pi w_k/h\nu_k)\left\{\sum_{j>i} \alpha_{ij}(\nu_k)[n_i - n_j(n_i/n_j)^* e^{-h\nu_k/kT}]\right. \\ &\qquad \left. - \sum_{j<i} \alpha_{ji}(\nu_k)[n_j - n_i(n_j/n_i)^* e^{-h\nu_k/kT}]\right\} \end{aligned} \tag{5-106}$$

We then can construct

$$(\partial \mathbf{n}/\partial J_k)_l = -\sum_m \mathscr{A}^{-1}_{lm} a_m \tag{5-107}$$

It is particularly instructive to consider the case where *only one* $(i \to j)$ transition can absorb at $\nu_k$, and all the other $\alpha_{lm}(\nu_k)$'s are zero.

*Exercise 5-9*: Show that, for the case just described, equation (5-108) is valid.

$$(\partial \mathbf{n}/\partial J_k)_l = (\mathscr{A}^{-1}_{lj} - \mathscr{A}^{-1}_{li})[4\pi w_k \alpha_{ij}(\nu_k)/h\nu_k][n_i - n_j(n_i/n_j)^* e^{-h\nu_k/kT}] \tag{5-108}$$

This result, besides being simple, shows clearly that, because $\mathscr{A}^{-1}$ must in general be a *full* matrix, a change in the radiation field at *any* frequency $\nu_k$ causes a change in the occupation numbers of *every* level $l$, *even if* $l$ *cannot absorb or emit photons of that frequency*. Of course the elements $\mathscr{A}^{-1}_{li}$ and $\mathscr{A}^{-1}_{lj}$ may be small, and the coupling weak, but the basic physical point remains true. Similar results for $(\partial/\partial n_e)$ and $(\partial/\partial T)$ are given in the references cited.

# 6

# *Solution of the Transfer Equation*

The analysis of actual stellar spectra requires the calculation of the emergent flux from a model atmosphere by a solution of the transfer equation. In this chapter we therefore turn attention to numerical techniques for solving transfer problems in terms of *differential equations*; we shall find that two extremely general, flexible, and powerful methods result when we formulate the transfer equation as a *two-point boundary-value problem* using *difference equations*. Many methods exist for solving the transfer equation in terms of *integral equations*, but these will not be discussed in depth in this book, as they are adequately described elsewhere (**18**, Chap. 8) and as they do not lend themselves so readily to the treatment of moving atmospheres (Chapter 14). In the present chapter we shall restrict attention to *static, one-dimensional plane-parallel atmospheres*; more general problems will be considered in Chapters 7 and 14. There are strong physical and mathematical motivations for using the techniques presented in § 6-3 (or their integral-equation analogues), which are specially designed to overcome certain difficulties that are characteristic of radiative transfer in optically thick media in the presence of scattering terms. We shall try to develop insight into these difficulties by considering other plausible, but useless, approaches to the problem in §§6-1 and 6-2.



## 6-1 Iteration: The Scattering Problem

One of the fundamental *physical* difficulties inherent in the solution of transfer problems is the existence of scattering terms, which decouple the radiation field from local sources and sinks, and involve global transport of photons over large distances in the atmosphere. It is through these terms that the presence of a free boundary makes itself felt even at great depth ($\tau_\nu \gg 1$) in an atmosphere, and causes large departures of the mean intensity $J_\nu$ from local values of the thermal source function $B_\nu$. For ease of discussion we consider a prototype source function that contains a thermal emission component and a coherent isotropic scattering term—i.e.,

$$S_\nu = (\kappa_\nu B_\nu + \sigma_\nu J_\nu)/(\kappa_\nu + \sigma_\nu) = (1 - \rho_\nu)B_\nu + \rho_\nu J_\nu \tag{2-39}$$

where $\rho_\nu \equiv \sigma_\nu/(\kappa_\nu + \sigma_\nu)$. The solution of the standard transfer equation

$$\mu(\partial I_\nu/\partial\tau_\nu) = I_\nu - S_\nu \tag{2-36}$$

can be written formally in terms of $J_\nu$ as [cf. Exercise (2-10)]

$$J_\nu(\tau_\nu) = \Lambda_{\tau_\nu}[S_\nu] = \Lambda_{\tau_\nu}[B_\nu] + \Lambda_{\tau_\nu}[\rho_\nu(J_\nu - B_\nu)]$$

If there were *no* scattering, $\rho_\nu \equiv 0$, then $J_\nu$ could be calculated, as a *quadrature*, from $B_\nu$; when $\rho_\nu \not\equiv 0$, we must solve an *integral equation* for $J_\nu$. One of the first methods that comes to mind to effect such a solution is *iteration*. As we know that $J_\nu \to B_\nu$ as $\tau_\nu \to \infty$, let us deal with $(J_\nu - B_\nu)$. Suppose $\rho_\nu$ were everywhere zero; then $(J_\nu - B_\nu)$ would equal $(\bar{B}_\nu - B_\nu)$ where $\bar{B}_\nu(\tau_\nu) \equiv \Lambda_{\tau_\nu}[B_\nu]$. If $\rho_\nu$ is not zero, we could regard this value as a first approximation and write

$$\begin{aligned}(J_\nu - B_\nu)^{(1)} &= (\bar{B}_\nu - B_\nu) + \Lambda_{\tau_\nu}[\rho_\nu(J_\nu - B_\nu)^0] \\ &= (\bar{B}_\nu - B_\nu) + \Lambda_{\tau_\nu}[\rho_\nu(\bar{B}_\nu - B_\nu)] \\ &= (\bar{B}_\nu - B_\nu) + \Delta^{(1)}\end{aligned} \tag{6-1}$$

Then by iteration, we find

$$(J_\nu - B_\nu)^{(n)} = (\bar{B}_\nu - B_\nu) + \sum_{i=1}^{n} \Delta^{(n)} \tag{6-2}$$

where $\Delta^{(n)} \equiv \Lambda_{\tau_\nu}[\rho_\nu \Delta^{(n-1)}]$. In practice we continue the iteration until some convergence criterion—e.g., $\|\Delta^{(n)}/(J_\nu - B_\nu)^{(n)}\| \leqslant \varepsilon$, where $\varepsilon \ll 1$—is satisfied. It is clear that, if $\|\rho_\nu\| \ll 1$, the iteration procedure of equation (6-2) can be expected to converge, for successive corrections $\Delta^{(n)}$ must be of order $\|\rho_\nu\|^n$ relative to $(J_\nu - B_\nu)$. If, however, $\|\rho_\nu\| \approx 1$ over a large depth of the atmosphere, the iteration method *fails*.

The circumstance just mentioned actually occurs in stellar atmospheres, and the thermal coupling parameter $\lambda_\nu \equiv 1 - \rho_\nu$ may be very small throughout a large part of the atmosphere. For example, in very hot stars the principal source of continuum opacity in the outer layers is electron scattering, and $\lambda_\nu$ may be of order $10^{-4}$ very deep into the atmosphere (until finally, as the density rises, free–free thermal absorption overwhelms the electron scattering). In cool stars of low metal abundance, the hydrogen is neutral in the upper atmosphere and free electrons are scarce, so Rayleigh scattering by H and $H_2$ dominates the $H^-$ opacity, and $\rho_\nu$ is nearly unity until great depth (at some point the hydrogen rather abruptly becomes excited and ionized, and $\lambda_\nu$ suddenly rises to unity). For lines, the corresponding thermal parameters may be very small, $\lambda_\nu \sim 10^{-8}$ (see Chapter 11).

The symptomatic behavior of the iteration method in these cases is that the solution *stabilizes*, and although successive iterations differ fractionally only by some small value, the $\Delta$'s are *monotonic*, and are *nearly equal* in iteration after iteration. In such cases, although the fractional change per iteration is $\varepsilon$ ($\varepsilon \ll 1$), there is no guarantee that, say, $1/\varepsilon$ more iterations may not actually be required to reach the final solution. The discussion thus far has been couched in terms of integral equations using the $\Lambda$-operator, but it should be stressed that the same difficulties would arise with a similar iterative solution of the transfer equation as a differential equation (we shall, in fact, refer to either procedure as "$\Lambda$-iteration" even when we do not actually employ the $\Lambda$-operator). The failure of $\Lambda$-iteration to converge is a point of crucial importance whose physical significance must be understood completely; to this end we may consider the following simplified analysis.

Suppose that the depth-variation of the Planck function can be represented with sufficient accuracy by a linear expansion

$$B_\nu(\tau_\nu) = a_\nu + b_\nu \tau_\nu \tag{6-3}$$

and that $\rho_\nu$ is constant with depth. The zero-order moment of the transfer equation can be written, using equations (2-71) and (2-39)

$$(\partial H_\nu/\partial \tau_\nu) = J_\nu - S_\nu = \lambda_\nu (J_\nu - B_\nu) \tag{6-4}$$

while the first-order moment gives

$$(\partial K_\nu/\partial \tau_\nu) = H_\nu \tag{6-5}$$

If we use the Eddington approximation $K_\nu = \frac{1}{3} J_\nu$ and substitute equation (6-5) for $H_\nu$ into equation (6-4) we obtain

$$\frac{1}{3}(\partial^2 J_\nu/\partial \tau_\nu{}^2) = \lambda_\nu (J_\nu - B_\nu) = \frac{1}{3}[\partial^2 (J_\nu - B_\nu)/\partial \tau_\nu{}^2] \tag{6-6}$$



where the second equality follows from the form of $B_\nu$ assumed in equation (6-3). The solution of equation (6-6) is

$$J_\nu - B_\nu = \alpha_\nu \exp[-(3\lambda_\nu)^{\frac{1}{2}} \tau_\nu] + \beta_\nu \exp[+(3\lambda_\nu)^{\frac{1}{2}} \tau_\nu] \tag{6-7}$$

As we demand that $J_\nu \to B_\nu$ as $\tau_\nu \to \infty$, we must have $\beta_\nu \equiv 0$. To evaluate $\alpha_\nu$ we make use of the boundary condition $J_\nu(0) = \sqrt{3}\, H_\nu(0) = (dJ_\nu/d\tau_\nu)_0/\sqrt{3}$ [the second equality following from equation (6-5) in the Eddington approximation]. We thus find from equation (6-7)

$$J_\nu(0) = a_\nu + \alpha_\nu = (dJ_\nu/d\tau)_0/\sqrt{3} = [b_\nu - \alpha_\nu (3\lambda_\nu)^{\frac{1}{2}}]/\sqrt{3} \tag{6-8}$$

Hence we obtain finally

$$J_\nu(\tau_\nu) = a_\nu + b_\nu \tau_\nu + (b_\nu - \sqrt{3}\, a_\nu) \exp[-(3\lambda_\nu)^{\frac{1}{2}} \tau_\nu]/[\sqrt{3} + (3\lambda_\nu)^{\frac{1}{2}}] \tag{6-9}$$

Equation (6-9) reveals the essential physics of the problem. First, it shows that $J_\nu$ may be markedly different from $B_\nu$ at the surface. For simplicity, consider an isothermal atmosphere—i.e., $b_\nu = 0$ and $B_\nu \equiv a_\nu$; then at $\tau_\nu = 0$, $J_\nu(0) = \lambda_\nu{}^{\frac{1}{2}} a_\nu/(1 + \lambda_\nu{}^{\frac{1}{2}}) = \lambda_\nu{}^{\frac{1}{2}} B_\nu/(1 + \lambda_\nu{}^{\frac{1}{2}})$. Thus when $\lambda_\nu \ll 1$, then $J_\nu$ is *much smaller* than $B_\nu$ at the boundary. Second, we see that this departure extends *deep* into the atmosphere because the slow decay of the exponential term implies that $J_\nu(\tau_\nu) \to B_\nu(\tau_\nu)$ *only at depths* $\tau_\nu \gtrsim (\lambda_\nu)^{-\frac{1}{2}}$; in view of the small values quoted above, these are *very large* depths indeed. When $J_\nu$ has approached $B_\nu$ arbitrarily closely, we say that the solution has *thermalized*; we therefore refer to $\lambda_\nu{}^{-\frac{1}{2}}$ as the *thermalization depth* (a concept that will be generalized in Chapters 7, 11, and 12).

We may obtain an intuitive understanding of the thermalization depth from the following physical argument. The parameter $\lambda_\nu = \kappa_\nu/(\kappa_\nu + \sigma_\nu)$ clearly is just the probability that a photon is destroyed (i.e., converted into thermal energy) per scattering event. To *assure* thermal destruction, the photon must be scattered about $n = 1/\lambda_\nu$ times. If the photon progresses through the atmosphere by a random-walk process, with mean free path $\Delta\tau$ (which must be approximately unity), then the total optical thickness through which it may pass without destruction is $n^{\frac{1}{2}}\, \Delta\tau = \Delta\tau \lambda_\nu{}^{-\frac{1}{2}} \approx \lambda_\nu{}^{-\frac{1}{2}}$. Photons emitted at greater depths are unlikely to escape without being thermalized (hence $J_\nu \to B_\nu$), while those emitted from shallower depths manage to escape and allow $J_\nu$ to fall below the thermal value (i.e., $B_\nu$).

We now can understand why $\Lambda$-iteration fails when we adopt $J_\nu = \bar{B}_\nu$ as an initial estimate. Each successive iteration can propagate information about the departure of $J_\nu$ from $\bar{B}_\nu$ only over an optical depth $\Delta\tau \approx 1$—i.e., a mean free path [recall that $E_1(\Delta\tau)$ falls off as $e^{-\Delta\tau}/\Delta\tau$ for $\Delta\tau \gg 1$]. Thus we must perform of order $\lambda_\nu{}^{-\frac{1}{2}}$ iterations to allow the effects of the boundary to make themselves felt in the solution to a thermalization depth. When

$\lambda_\nu \ll 1$ such a procedure becomes computationally prohibitive, and we conclude that any useful method must account for the scattering terms in the source function *from the outset* and provide a *direct* solution for such terms.

## 6-2 Eigenvalue Methods

A characteristic *mathematical* difficulty that emerges in treating the transfer equation as a differential equation arises from the nature of the boundary conditions. Suppose we use the method of discrete ordinates, replace the angular integral for $J_\nu$ by a quadrature sum, and attempt to integrate numerically the system

$$\mu_i(dI_i/d\tau) = I_i - \frac{1}{2}\rho \sum_{j=-n}^{n} a_j I_j - (1-\rho)B, \quad (i = \pm 1, \ldots, \pm n) \quad (6\text{-}10)$$

To effect the integration, we require starting values for $I_i$, for all values of $i$; these are fixed by the boundary conditions. As described in Chapter 2, the boundary conditions fall into *two* groups, namely $I_i(0) = 0$, $(i = -1, \ldots, -n)$, for incoming rays on the range $-1 \leqslant \mu_i \leqslant 0$ and $I_i(\tau_{\max}) = g(\mu_i)$ [e.g., $g(\mu) \equiv B_\nu$], $(i = 1, \ldots, n)$, for outgoing rays on the range $0 \leqslant \mu_i \leqslant 1$. Here $\tau_{\max}$ refers to the deepest point actually treated in a semi-infinite atmosphere. The problem is this: suppose we wish to start the integration at $\tau = 0$, and proceed step-by-step inward; we cannot, for we do not know the values of $I_i(0)$. Similarly at $\tau_{\max}$ we lack values for $I_{-i}(\tau_{\max})$.

Thus we face an *eigenvalue problem* of order $n$. We could, for example, guess a set of values for $I_{-i}(\tau_{\max})$ and use these to integrate toward the surface. When the integration reaches the surface, we would in general find $I_{-i}(0) \neq 0$. In principle, we could then adjust the values of $I_{-i}(\tau_{\max})$, and by successive trials find those values that forced $I_{-i}(0) = 0$. In practice, however, this method is strongly unstable and can work only if $\tau_{\max}$ is not very large. We can see this as follows. As we know from the grey problem, the discrete ordinate method leads to exponential solutions of the form $\exp(\pm k\tau)$ where the $k$'s are of order $1/\mu$. In cases where the coefficients (such as $\rho_\nu$) are depth-variable, the solution no longer consists of pure exponentials, but, nevertheless, still has an exponential character, perhaps $f(\tau)\exp(\pm k\tau)$ where $f$ is a weak function of $\tau$. In a semi-infinite atmosphere we must suppress the ascending exponentials. For the grey problem this can be done explicitly, for we have an analytical form with which to work. But in the nongrey variable-coefficient case, the solution is known only numerically, and unless *exactly* the right choice of starting values is made, it contains both the ascending and descending exponentials. Therefore, in general the terms in $\exp(k\tau)$ will be present; these are called *parasites*, and they increase at a rate of order $\exp(2k\tau)$ relative to the true solution. Thus if our starting values are wrong by an error $\varepsilon$, the parasite will be of order $\varepsilon \exp(2k\tau_{\max}) \sim \varepsilon\, 10^{k\tau_{\max}}$ compared to the true solution at the other boundary, and it is obvious that, unless our initial choice is *very* good ($\varepsilon \ll 1$), the parasite will swamp the true solution, which will then be lost. In fact, to retain any vestige of the real solution, we must employ $n \approx k\tau_{\max}$ significant figures. If several angle-quadrature points are used, some $\mu_i \ll 1$ and hence some $k \gg 1$, so even with a moderate $\tau_{\max} \approx 10$ we will lose the solution on typical computers. At $\tau \approx 1$ in the continuum, $\tau_{\max}$ may be $\sim 10^3$ to $10^4$ in the lines, which shows the hopelessness of this approach. In summary, the mathematical structure of the problem requires that we employ a method that accounts explicitly for the two-point nature of the boundary conditions from the outset. We now turn to a discussion of such methods.



*Exercise 6-1*: (a) Solve equation (6-10) with $\rho \equiv 0$, $B = \text{const}$, for $I_\pm$ with $\mu_\pm = \pm\frac{1}{2}$. Show that $d^2J/d\tau^2 = 4(J - B)$ and write exact solutions for $J$, $I_+$, and $I_-$, calculating constants of integration from boundary conditions. Suppose one had chosen $I_+(\tau_{\max}) = I_-(\tau_{\max}) \equiv B$; evaluate the (false) solution and show that the error $\varepsilon = B\exp(-2\tau_{\max})$ at the lower boundary amplifies to $\varepsilon = B$ at the surface. (b) Generalize the discussion to the case where $\rho \neq 0$ (but constant).

## 6-3 The Transfer Equation as a Two-Point Boundary Value Problem

In this section we shall derive two very general, flexible, and powerful approaches for solving transfer problems. These approaches result from writing the transfer equation as a second-order differential equation subject to two-point boundary conditions. Most of the basic ideas were presented in an important paper by Feautrier (**209**). These methods have proven to be stable and easy to implement; each offers advantages in complementary ranges of the parameters that set the scale of the computational effort to solve a given problem.

### SECOND-ORDER FORM OF THE EQUATION OF TRANSFER

In plane-parallel geometry we may write two equations governing the outgoing and incoming radiation field at $\pm\mu$:

$$\pm\mu[\partial I(z, \pm\mu, \nu)/\partial z] = \chi(z, \nu)[S(z, \nu) - I(z, \pm\mu, \nu)] \quad (6\text{-}11)$$

where we restrict $\mu$ to the half-range $0 \leqslant \mu \leqslant 1$. We now define symmetric and antisymmetric averages

$$u(z, \mu, \nu) \equiv \frac{1}{2}[I(z, \mu, \nu) + I(z, -\mu, \nu)] \tag{6-12}$$

and

$$v(z, \mu, \nu) \equiv \frac{1}{2}[I(z, \mu, \nu) - I(z, -\mu, \nu)] \tag{6-13}$$

which have, respectively, a mean-intensity-like and a flux-like character. In terms of $u$ and $v$ we can construct a system of two first-order equations by adding the two equations (6-11) to obtain

$$\mu[\partial v(z, \mu, \nu)/\partial z] = \chi(z, \nu)[S(z, \nu) - u(z, \mu, \nu)] \tag{6-14}$$

and subtracting them to obtain

$$\mu[\partial u(z, \mu, \nu)/\partial z] = -\chi(z, \nu)v(z, \mu, \nu) \tag{6-15}$$

Then substituting equation (6-15) into (6-14) we can eliminate $v$ and obtain a single second-order system for $u$:

$$\frac{\mu^2}{\chi(z, \nu)}\frac{\partial}{\partial z}\left[\frac{1}{\chi(z, \nu)}\frac{\partial u(z, \mu, \nu)}{\partial z}\right] = u(z, \mu, \nu) - S(z, \nu) \tag{6-16}$$

or, defining $d\tau_\nu \equiv d\tau(z, \nu) = -\chi(z, \nu)\, dz$ and abbreviating the notation,

$$\mu^2(\partial^2 u_{\mu\nu}/\partial\tau_\nu{}^2) = u_{\mu\nu} - S_\nu \tag{6-17}$$

In writing equation (6-15) we have assumed that $S$ is symmetric in $\mu$; this will be true for most of the source functions we shall consider—e.g., those of the form

$$S_\nu = \alpha_\nu \int \phi_{\nu'} J_{\nu'}\, d\nu' + \beta_\nu \tag{6-18}$$

or

$$S_\nu = \alpha_\nu \int R(\nu', \nu) J_{\nu'}\, d\nu' + \beta_\nu \tag{6-19}$$

but may not be true if the redistribution is angle-dependent [in which case other techniques are required, cf. (**460**)] or if there are motions in the atmosphere (see §14-1). In equations (6-18) and (6-19) the $\alpha$'s essentially stand for scattering coefficients divided by the total opacity and the $\beta$'s represent thermal terms. It must be stressed that these choices of $S_\nu$ are purely *illustrative*, in the sense that we shall later (cf. §§7-2 and 7-5) find similar-looking terms that involve the radiation field over the *entire spectrum* (imposed by a radiative equilibrium constraint) or for the *entire transition array* for a multilevel model atom. The analysis given below still applies in such cases.



Note that in contrast to the moment equations, which do not close, equation (6-17) [first derived by Feautrier (**209**)] yields exact closure of the system in terms of the angle-dependent symmetric average $u_{\mu\nu}$. We shall see below that it is sometimes advantageous to follow an intermediate course and to use an approximate closure of the moment equations in terms of variable Eddington factors.

### BOUNDARY CONDITIONS

Equation (6-17) must be solved subject to boundary conditions at $\tau = 0$ and at $\tau = \tau_{\max}$ [which denotes the thickness (or half-thickness) of a finite slab, or a great depth where the diffusion approximation applies for a semi-infinite atmosphere]. At $\tau = 0$, $I(0, -\mu, \nu) \equiv 0$ which implies that $v_{\mu\nu}(0) \equiv u_{\mu\nu}(0)$ so that

$$\mu(\partial u_{\mu\nu}/\partial\tau_\nu)_0 = u_{\mu\nu}(0) \tag{6-20}$$

At $\tau = \tau_{\max}$, we specify $I(\tau_{\max}, +\mu, \nu) = I_+(\mu, \nu)$, and write $v_{\mu\nu}(\tau_{\max}) = I_+(\mu, \nu) - u_{\mu\nu}(\tau_{\max})$ so that

$$\mu(\partial u_{\mu\nu}/\partial\tau_\nu)_{\tau_{\max}} = I_+(\mu, \nu) - u_{\mu\nu}(\tau_{\max}) \tag{6-21}$$

If the diffusion approximation is valid at $\tau_{\max}$, then

$$I(\tau_{\max}, \mu, \nu) = B_\nu(\tau_{\max}) + \mu\left(\frac{1}{\chi_\nu}\left|\frac{\partial B_\nu}{\partial z}\right|\right)_{\tau_{\max}} \tag{6-22}$$

so that $u_{\mu\nu}(\tau_{\max}) = B_\nu(\tau_{\max})$, $v_{\mu\nu}(\tau_{\max}) = \mu(\chi_\nu{}^{-1}|\partial B_\nu/\partial z|)_{\tau_{\max}}$ and

$$\mu\left.\frac{\partial u_{\mu\nu}}{\partial\tau_\nu}\right|_{\tau_{\max}} = \mu\left(\frac{1}{\chi_\nu}\left|\frac{\partial B_\nu}{\partial z}\right|\right)_{\tau_{\max}} \tag{6-23}$$

*Exercise 6-2*: (a) Generalize equation (6-20) when $I(0, -\mu, \nu) \neq 0$. (b) Show that for a *symmetric slab* (infinite in $x$ and $y$), of finite thickness (in $z$) $\tau_{\max}$, the lower boundary condition can be written at $\tau = \frac{1}{2}\tau_{\max}$ as $(\partial u_{\mu\nu}/\partial\tau_\nu) = 0$. This implies that we need consider only half the slab: $0 \leqslant \tau \leqslant \frac{1}{2}\tau_{\max}$.

### DIFFERENCE-EQUATION REPRESENTATION

We now convert the differential equation (6-17) into a set of *difference equations* by discretization of all variables. Thus we choose a set of depth points $\{\tau_d\}$, $(d = 1, \ldots, D)$ with $\tau_1 < \tau_2 < \cdots < \tau_D$; a set of angle points $\{\mu_m\}$, $(m = 1, \ldots, M)$; and a set of frequency points $\{\nu_n\}$, $(n = 1, \ldots, N)$. For any variable $g$, we write $g(z_d, \mu_m, \nu_n) = g_{dmn}$. We replace integrals by

quadrature sums—e.g., for equation (6-18) we write

$$S_{dn} = \alpha_{dn} \sum_{n=1}^{N} a_n \phi_{dn} \sum_{m=1}^{M} b_m u_{dmn} + \beta_{dn} \tag{6-24a}$$

Further, we group angles and frequencies into a single serial set of values subscripted $i$ such that $(\mu_i, \nu_i) = (\mu_m, \nu_n)$ at $i = m + (n-1)M$, and hence reduce (6-24a) to

$$S_{di} = \alpha_{di} \sum_{i'=1}^{I} w_{i'} \phi_{di'} u_{di'} + \beta_{di} \quad (i = 1, \ldots, I) \tag{6-24b}$$

Similarly, for equation (6-19) we have

$$S_{di} = \alpha_{di} \sum_{i'=1}^{I} \mathscr{R}_{d,i',i} u_{di'} + \beta_{di} \quad (i = 1, \ldots, I) \tag{6-25}$$

Note in passing that these source functions are independent of angle, and hence this description contains redundant information (which can be removed when we introduce variable Eddington factors). Equation (6-24b) has an additional redundancy because the scattering integral is independent of $\nu$ (or of $i$); we shall exploit this later in Rybicki's method of solving the equations.

Further, we replace derivatives by difference formulae, and write, e.g.,

$$(dX/d\tau)_{d+\frac{1}{2}} \approx (\Delta X_{d+\frac{1}{2}}/\Delta\tau_{d+\frac{1}{2}}) = (X_{d+1} - X_d)/(\tau_{d+1} - \tau_d) \tag{6-26}$$

and

$$(d^2X/d\tau^2)_d \approx \left[(dX/d\tau)_{d+\frac{1}{2}} - (dX/d\tau)_{d-\frac{1}{2}}\right] \Big/ \left[\frac{1}{2}(\Delta\tau_{d+\frac{1}{2}} + \Delta\tau_{d-\frac{1}{2}})\right] \tag{6-27}$$

thus, defining

$$\Delta\tau_{d\pm\frac{1}{2},i} \equiv \frac{1}{2}(\chi_{d\pm1,i} + \chi_{di})|z_{d\pm1} - z_d| \tag{6-28}$$

and

$$\Delta\tau_{d,i} \equiv \frac{1}{2}(\Delta\tau_{d-\frac{1}{2},i} + \Delta\tau_{d+\frac{1}{2},i}) \tag{6-29}$$

we rewrite equation (6-17) as

$$\left(\frac{\mu_i^2}{\Delta\tau_{d-\frac{1}{2},i}\Delta\tau_{d,i}}\right) u_{d-1,i} - \frac{\mu_i^2}{\Delta\tau_{d,i}}\left(\frac{1}{\Delta\tau_{d-\frac{1}{2},i}} + \frac{1}{\Delta\tau_{d+\frac{1}{2},i}}\right) u_{d,i} + \left(\frac{\mu_i^2}{\Delta\tau_{d,i}\Delta\tau_{d+\frac{1}{2},i}}\right) u_{d+1,i} = u_{di} - S_{di} \quad \begin{matrix}(i = 1, \ldots, I) \\ (d = 2, \ldots, D-1)\end{matrix} \tag{6-30}$$



where $S_{di}$ has the form of equation (6-24) or (6-25), and indeed can be generalized still further (to include, e.g., the *entire* spectrum; cf. §§7-2 and 7-5). As indicated, there is one such equation for each angle–frequency point $i$, at each of $D-2$ depth points.

If we now define the vector $\mathbf{u}_d$, of dimension $I$, to consist of the angle–frequency components at depth-point $d$—i.e., $(\mathbf{u}_d)_i = u_{di}$—then equation (6-30) can be written as a matrix equation

$$-\mathbf{A}_d\mathbf{u}_{d-1} + \mathbf{B}_d\mathbf{u}_d - \mathbf{C}_d\mathbf{u}_{d+1} = \mathbf{L}_d \tag{6-31}$$

The $(I \times I)$ matrices $\mathbf{A}_d$ and $\mathbf{C}_d$ are *diagonal* and contain the finite-difference representation of the differential operator. $\mathbf{B}_d$ is a *full* matrix that has the differential operator down the diagonal plus off-diagonal terms that come from the quadrature sum representing the scattering integrals in equations (6-24) and (6-25). $\mathbf{L}_d$ is a vector containing the thermal source terms. More accurate difference representations than equation (6-30) may be written using spline colocation (**374**), (**442**) or Hermite integration formulae (**34**), but these do not change the general form of equation (6-31) (though $\mathbf{A}_d$ and $\mathbf{C}_d$ may become full).

To complete the system, we use the boundary conditions. At the surface we could write

$$\mu_i(u_{2i} - u_{1i})/\Delta\tau_{\frac{3}{2},i} = u_{1,i} \tag{6-32}$$

which is only of first-order accuracy; second-order accuracy can be obtained (**30**) from the Taylor's expansion $u_2 = u_1 + \Delta\tau_{\frac{3}{2}}(du/d\tau)_1 + \frac{1}{2}\Delta\tau_{\frac{3}{2}}^2(d^2u/d\tau^2)_1$, which, using equations (6-17) and (6-20), yields

$$\mu_i(u_{2i} - u_{1i})/\Delta\tau_{\frac{3}{2},i} = u_{1i} + \left(\frac{1}{2}\Delta\tau_{\frac{3}{2},i}\Big/\mu_i\right)(u_{1i} - S_{1i}) \tag{6-33}$$

or, in matrix form

$$\mathbf{B}_1\mathbf{u}_1 - \mathbf{C}_1\mathbf{u}_2 = \mathbf{L}_1 \tag{6-34}$$

Similarly, equation (6-21) at the lower boundary becomes

$$\mu_i(u_{Di} - u_{D-1,i})/\Delta\tau_{D-\frac{1}{2},i} = I_{Di}^{+} - u_{Di} - \left(\frac{1}{2}\Delta\tau_{D-\frac{1}{2},i}\Big/\mu_i\right)(u_{Di} - S_{Di}) \tag{6-35}$$

which, in matrix form is

$$-\mathbf{A}_D\mathbf{u}_{D-1} + \mathbf{B}_D\mathbf{u}_D = \mathbf{L}_D \tag{6-36}$$

Note that $\mathbf{A}_1 \equiv 0$ and $\mathbf{C}_D \equiv 0$.

*Exercise 6-3*: Derive equations (6-33) and (6-35), specialize the latter to the diffusion approximation using equation (6-23).

THE FEAUTRIER SOLUTION

The set of equations (6-31), (6-34), and (6-36) have the overall structure

$$\begin{pmatrix} \mathbf{B}_1 & -\mathbf{C}_1 & & & & \\ -\mathbf{A}_2 & \mathbf{B}_2 & -\mathbf{C}_2 & & & \\ 0 & -\mathbf{A}_3 & \mathbf{B}_3 & -\mathbf{C}_3 & & \\ & & & \ddots & & \\ & & & -\mathbf{A}_{D-1} & \mathbf{B}_{D-1} & -\mathbf{C}_{D-1} \\ & & & & -\mathbf{A}_D & \mathbf{B}_D \end{pmatrix} \begin{pmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \\ \mathbf{u}_3 \\ \vdots \\ \mathbf{u}_{D-1} \\ \mathbf{u}_D \end{pmatrix} = \begin{pmatrix} \mathbf{L}_1 \\ \mathbf{L}_2 \\ \mathbf{L}_3 \\ \vdots \\ \mathbf{L}_{D-1} \\ \mathbf{L}_D \end{pmatrix} \tag{6-37}$$

Each element indicated is either an $(I \times I)$ matrix or a vector of length $I$; the grand matrix has a *block tridiagonal* structure, and the solution proceeds by an efficient forward-elimination and back-substitution procedure (**209**). In this scheme we in effect express each $\mathbf{u}_d$ in terms of $\mathbf{u}_{d+1}$ and substitute into the following equation. Thus from equation (6-34) we can write

$$\mathbf{u}_1 = \mathbf{B}_1^{-1}\mathbf{C}_1\mathbf{u}_2 + \mathbf{B}_1^{-1}\mathbf{L}_1 \equiv \mathbf{D}_1\mathbf{u}_2 + \mathbf{v}_1 \tag{6-38}$$

Substituting equation (6-38) into equation (6-31) for $d = 2$ yields $\mathbf{u}_2 = \mathbf{D}_2\mathbf{u}_3 + \mathbf{v}_2$ where $\mathbf{D}_2 = (\mathbf{B}_2 - \mathbf{A}_2\mathbf{D}_1)^{-1}\mathbf{C}_2$

and $$\mathbf{v}_2 = (\mathbf{B}_2 - \mathbf{A}_2\mathbf{D}_1)^{-1}(\mathbf{L}_2 + \mathbf{A}_2\mathbf{v}_1)$$

We therefore have in general

$$\mathbf{u}_d = \mathbf{D}_d\mathbf{u}_{d+1} + \mathbf{v}_d \tag{6-39}$$

where $$\mathbf{D}_d \equiv (\mathbf{B}_d - \mathbf{A}_d\mathbf{D}_{d-1})^{-1}\mathbf{C}_d \tag{6-40}$$

and $$\mathbf{v}_d \equiv (\mathbf{B}_d - \mathbf{A}_d\mathbf{D}_{d-1})^{-1}(\mathbf{L}_d + \mathbf{A}_d\mathbf{v}_{d-1}) \tag{6-41}$$

for $d = 1, \ldots, D$. Starting at $d = 1$, we compute successive values for $\mathbf{D}_d$ and $\mathbf{v}_d$ through $d = D - 1$. At the last point, $d = D$, $\mathbf{C}_D \equiv 0$, hence $\mathbf{D}_D \equiv 0$, and $\mathbf{u}_D \equiv \mathbf{v}_D$ [which still follows from equation (6-41)]. Having found $\mathbf{u}_D$ we then perform successive back-substitutions into equation (6-39) to find $\mathbf{u}_d$, $(d = D - 1, \ldots, 2, 1)$. Having found $u_{dmn}$, we may then evaluate $J_{dn} \equiv \sum_{m=1}^{M} b_m u_{dmn}$, and the source function, which involves frequency integrals of $J_\nu$: e.g., $S_{dn} = \alpha_{dn} \sum_{n'} w_{n'}\phi_{dn'}J_{dn'} + \beta_{dn}$.

The forward–backward sweep described above accounts explicitly for scattering terms and the two-point boundary conditions. Feautrier's method has proven itself to be very stable, and has many desirable properties. Note, for example, that at depth the system tends to become diagonal (the terms in $1/\Delta\tau^2 \to 0$) and hence $\mathbf{J}_d \to \mathbf{S}_d$, as expected; in fact, we find $\mathbf{J} \to \mathbf{S} + \mu^2(d^2\mathbf{S}/d\tau^2)$, which recovers the diffusion approximation automatically. The depth-discretization is commonly taken to give equally-spaced steps in $\log\tau$, usually with 5 or 6 steps per decade of $\tau$; such a choice has the advantage that at different frequencies with widely differing opacities (e.g., a line-core vs. nearby continuum) one has a reasonable distribution of depth-points.



We can estimate the amount of computing time required in a given problem by counting the number of multiplications needed to solve the system; the solution of a linear system of order $n$ requires $0(n^3)$ operations, so the time required by Feautrier's method is $T_F = cDI^3 = cDM^3N^3$ where $D$ is the number of depth-points, $M$ the number of angle-points, and $N$ the number of frequency points. It is clear that one pays a penalty for any unnecessary redundancy in the angle–frequency information, and that the representation of these variables must be economized as much as possible. If we have a *coherent* scattering problem, $N = 1$, $M$ is generally small, and Feautrier's method is *optimum*. However, in other problems the number of frequencies can be large because we must satisfy the constraint of radiative equilibrium, or statistical equilibrium in several transitions; but the angular information is essentially unnecessary because only $J_\nu$, not $u_{\mu\nu}$, enters these constraints. We therefore eliminate the angular information by introducing variable Eddington factors $f_\nu \equiv K_\nu/J_\nu$ (**44**).

By integration of equation (6-17) over $\mu$ we obtain

$$\partial^2(f_\nu J_\nu)/\partial\tau_\nu^2 = J_\nu - S_\nu \tag{6-42}$$

and the boundary conditions yield

$$[\partial(f_\nu J_\nu)/\partial\tau_\nu]_0 = h_\nu J_\nu(0) \tag{6-43}$$

and $$\left.\frac{\partial(f_\nu J_\nu)}{\partial\tau_\nu}\right|_{\tau_{\max}} = \frac{1}{3}\left(\frac{1}{\chi_\nu}\left|\frac{\partial B_\nu}{\partial z}\right|\right)_{\tau_{\max}} \tag{6-44}$$

where $h_\nu \equiv H_\nu(0)/J_\nu(0)$. Equations (6-42) through (6-44) may be differenced in the same way as the angle-dependent equations, but the solution of this system requires a time of only $T_\nu = cDN^3$, which represents a considerable saving. To solve these equations we must know the depth-variation of $f_\nu$ at all frequencies. We proceed as follows. (a) From any *given* $S_\nu$ (e.g., $S_\nu \equiv B_\nu$ as a first estimate) we can solve equation (6-17) for $u_{\mu\nu}$ *one angle and frequency at a time*. In matrix form we have $\mathbf{T}_i\mathbf{u}_i = \mathbf{S}_i$ where $\mathbf{T}$ is tridiagonal, and $\mathbf{u}_i$ and $\mathbf{S}_i$ represent the *depth-variation* of $u_{di}$ and $S_{di}$ respectively. Solution of a single tridiagonal system of order $n$ requires $0(n)$ operations, so the time required to evaluate the full angle-dependent radiation field for

*given* $S_\nu$ is $T_\mu = c'DMN$. (b) Given $u_{dmn}$ we then calculate

$$f_{dn} = \sum_m b_m \mu_m{}^2 u_{dmn} \Big/ \sum b_m u_{dmn}$$

and $h_n = \sum_m b_m \mu_m u_{1mn} / \sum b_m u_{1mn}$. Note that even if the radiation field is known only with modest accuracy, the Eddington factor may be determined with substantially better precision (e.g., if $u$ is in error by a scale-factor, $f$ is still correct). (c) Now, given $f_{dn}$, we solve equations (6-42)–(6-44) for $J_\nu$ using *explicit expressions of the form of equations (6-24) and (6-25) for* $S_\nu$ (written in terms of $J_\nu$). We then re-evaluate $S_\nu$ using the new values for $J_\nu$. (d) Because $S_\nu$ found in step (c) differs from that used in step (a), we *iterate* steps (a)–(c) to convergence. If $L$ is the number of iterations the total computing time is $T_E = L(cDN^3 + c'DMN) \ll cDM^3N^3$ for moderate $L$. Experience with this method for a very wide variety of physical regimes in stellar atmospheres has always shown extremely rapid convergence ($L$ usually is 3 or 4), and substantial economies (about a factor of ten) are achieved. Finally, we note that additional equations can be added to the transfer equations at each depth-point $d$; these arise from other physical constraints—e.g., statistical, hydrostatic, or radiative equilibrium (see §7-5). The basic form of equation (6-31) remains unaltered because these constraints involve information only at one or two depth-points at a time. Thus if we have $C$ constraints the total computing time becomes $T_E = L[cD(N + C)^3 + c'DMN]$; this result bears on the question of whether it is advantageous to use Feautrier's solution or Rybicki's solution, which we shall discuss next.

*Exercise 6-4*: This exercise requires access to a digital computer (of small capacity). (a) Write a computer program to perform the formal solution of the transfer equation with a *given* $S_\nu$ for $u_{\mu\nu}$, one angle at a time as described above, and to evaluate the variable Eddington factors at all depths. Use equally-spaced steps in $\Delta \log \tau$ starting at $\tau = 10^{-3}$, up to $\tau = 10$ (5 or 6 steps per decade), and use a double-Gauss angle quadrature (**4**, 921); experiment with the number of angle-points $M$ to examine the sensitivity of the Eddington factors to the quadrature. (b) Write a computer program to solve equations (6-42) through (6-44) with *given* Eddington factors, assuming coherent scattering—i.e., $S_\nu = \alpha J_\nu + \beta$. Integrate the two programs and study the convergence of the iteration process in cases with $\alpha = (1 - \varepsilon)$, $\beta = \varepsilon$, $\varepsilon \ll 1$, starting with $J_\nu \equiv 1$, for $\varepsilon = 0.1$, $0.01, 10^{-4}$.

### THE RYBICKI SOLUTION

As we have seen above, the Feautrier solution organizes the calculation in such a way as to group all frequency information together at a given depth, and to solve depth-by-depth; in that method we may treat a fully *frequency-dependent source function* [e.g., equation (6-25)] with *partial redistribution*, but the computing time scales as the cube of number of frequency points. In a beautiful paper, Rybicki (**543**) pointed out that, in the most commonly considered case of *complete redistribution*, much of this frequency-dependent information is redundant, for to specify the source function [equation (6-24)] we need only the single quantity $\bar{J} \equiv \int \phi_\nu J_\nu \, d\nu$. In a penetrating analysis, Rybicki showed how the solution in this case could be reorganized to yield a system of as great power and generality as the original Feautrier method but with very favorable computing time requirements.



Instead of describing the frequency variation of the radiation field at a given depth, let us instead reverse the grouping and work with vectors that describe the depth-variation at a given frequency. That is, we now define

$$\mathbf{u}_i \equiv (u_{1i}, u_{2i}, \ldots, u_{Di})^T \tag{6-45}$$

where $i$ denotes a particular angle–frequency point. Similarly let

$$\bar{\mathbf{J}} \equiv (\bar{J}_1, \bar{J}_2, \ldots, \bar{J}_D)^T \tag{6-46}$$

Then at angle–frequency point $i$, equations (6-30), (6-33), and (6-35) yield a system

$$\mathbf{T}_i \mathbf{u}_i + \mathbf{U}_i \bar{\mathbf{J}} = \mathbf{K}_i, \quad (i = 1, \ldots, I) \tag{6-47}$$

where $\mathbf{T}_i$ is a $(D \times D)$ *tridiagonal* matrix representing the differential operator at frequency $i$, $\mathbf{U}_i$ is a *diagonal* matrix containing the depth-variation of the scattering coefficient [$\alpha_{di}$ in equation (6-24)], and $\mathbf{K}_i$ is a vector that contains the depth-variation of the thermal term at frequency $i$. We have one set of equations (6-47) for each angle–frequency point. In addition we have $D$ equations that define $\bar{J}_d$, namely

$$\sum_{i'=1}^{I} w_{i'} \phi_{di'} u_{di'} - \bar{J}_d = 0, \quad (d = 1, \ldots, D) \tag{6-48}$$

Considering the grand matrix composed of all angle-frequencies and all depths, we have for the overall structure

$$\begin{pmatrix} \mathbf{T}_1 & & & \mathbf{U}_1 \\ & \mathbf{T}_2 & & \mathbf{U}_2 \\ & & \ddots & \vdots \\ & & \mathbf{T}_I & \mathbf{U}_I \\ \mathbf{V}_1 & \mathbf{V}_2 & \cdots \; \mathbf{V}_I & \mathbf{E} \end{pmatrix} \begin{pmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \\ \vdots \\ \mathbf{u}_I \\ \bar{\mathbf{J}} \end{pmatrix} = \begin{pmatrix} \mathbf{K}_1 \\ \mathbf{K}_2 \\ \vdots \\ \mathbf{K}_I \\ \mathbf{P} \end{pmatrix} \tag{6-49}$$

where the **V**'s are $(D \times D)$ *diagonal* matrices containing the depth variation of the quadrature weights and profiles in equation (6-48). **E** would be the negative identity matrix and **P** would be void for equation (6-48) but, as we shall see in §7-2, more general systems of the form of equation (6-49) arise in the computation of LTE model atmospheres. Comparison of equations (6-49) and (6-37) reveals that, in essence, the inner and outer structure of the equations has been interchanged.

The solution of system (6-49) is very efficient. We reduce each "row" of the grand matrix by solving for

$$\mathbf{u}_i = (\mathbf{T}_i^{-1}\mathbf{K}_i) - (\mathbf{T}_i^{-1}\mathbf{U}_i)\bar{\mathbf{J}} \quad (i = 1, \ldots, I) \tag{6-50}$$

Then, multiplying by $\mathbf{V}_i$ and subtracting from the last "row" we cancel the "element" in the $i$th column to zero. Carrying out this procedure for all values $i$ we obtain a final system for $\bar{\mathbf{J}}$, namely $\mathbf{W}\bar{\mathbf{J}} = \mathbf{Q}$ where the full $(D \times D)$ matrix **W** is

$$\mathbf{W} \equiv \mathbf{E} - \sum_{i=1}^{I} \mathbf{V}_i\mathbf{T}_i^{-1}\mathbf{U}_i \tag{6-51}$$

and the vector **Q** is

$$\mathbf{Q} = \mathbf{P} - \sum_{i=1}^{I} \mathbf{V}_i\mathbf{T}_i^{-1}\mathbf{K}_i \tag{6-52}$$

The final system for $\bar{\mathbf{J}}$ is solved; this yields sufficient information to calculate **S** (the run of the source function with depth). If desired, the full angle–frequency variation of the radiation field may be reconstructed using the already-available quantities $\mathbf{T}_i^{-1}\mathbf{K}_i$ and $\mathbf{T}_i^{-1}\mathbf{U}_i$ in equation (6-50).

The solution of the $I$ tridiagonal systems in equation (6-50) requires $0(D^2I) = 0(D^2MN)$ operations, and solution of the final system requires $0(D^3)$ operations, so the overall computing time becomes $T_R = cD^2MN + c'D^3$ (NOTE: these $c$'s are not numerically equal to those in the formulae for $T_F$, $T_E$, etc.). Unlike the Feautrier system, in which the computing time scales as the *cube* of the number of angle–frequency points (i.e., $M^3N^3$), Rybicki's method is *linear* in $MN$. It is obvious that Rybicki's method is vastly more economical than Feautrier's (even with variable Eddington factors) when a large number of frequency-points is required. Recall, however, that Rybicki's method works *only* if $S_\nu$ can be written in terms of a *single* quantity $\bar{J}$ in the scattering integral, while Feautrier's method works for *general* scattering integrals. In principle one could use variable Eddington factors with Rybicki's method, but the advantage gained would likely be small (if any) because iterations would then be required. It should also be emphasized that Rybicki's method is exactly equivalent to the integral equation approach in which one writes $\mathbf{u}_i = \mathbf{\Lambda}_i\bar{\mathbf{J}} + \mathbf{M}_i$, where the $\mathbf{\Lambda}$ matrix is generated by analytical integration of the kernel function against a set of basis functions representing $\bar{\mathbf{J}}$. In fact, $\mathbf{T}_i^{-1}$ *is* the $\mathbf{\Lambda}_i$ matrix, and inversion



of $\mathbf{T}_i$ is markedly less costly than any other approach for generating $\mathbf{\Lambda}$ (**34**); put another way, one *may* use integral-equation techniques if one wishes, but one should do so by means of Rybicki's method for generating $\mathbf{\Lambda}_i$.

*Exercise 6-5*: Using a digital computer, write a program to solve the transfer equation by Rybicki's method for a coherent scattering source function $S_\nu = \alpha J_\nu + \beta$ for the same values of $\varepsilon$ as were used in Exercise 6-4. Note that the Rybicki method does not show its advantage here because only one frequency-point is involved.

Finally, let us mention the effects of constraints in Rybicki's solution. For each constraint that introduces essentially new information into the problem, one requires an additional new variable similar to $\bar{\mathbf{J}}$, along with its defining equations. For example, in a *multiplet* problem (see §12-3) one requires a $\bar{\mathbf{J}}$ for each independent transition, and in problems where one has introduced the full set of statistical equilibrium equations by linearization a new variable is required for each level of the model atom or every line in the transition array (see §12-4). If we have a total of $C$ variables describing the constraints, then each **U** matrix must consist of $C$ diagonal $(D \times D)$ matrices side by side while each **V** matrix consists of $C$ diagonal $(D \times D)$ matrices stacked into a column, and **E** becomes a matrix of dimension $(CD \times CD)$. In this case the computing time for a direct solution becomes

$$T_R = c(D^2 \cdot M \cdot N \cdot C) + c'(DC)^3$$

for $C \gg 1$ this value exceeds the corresponding value for $T_F$, and at first sight Feautrier's method looks more attractive for dealing with systems involving many constraints (which is why we shall apply it in §7-5 for non-LTE model construction). Nevertheless, for *statistical equilibrium* calculations, Rybicki's method has been applied successfully even for large values of $C$ by using an iterative solution of the overall system (cf. §12-4).

### COMPUTATION OF THE FLUX

To compare with observations, we must calculate the emergent flux. This may be done in a variety of ways. If Feautrier's method is used with variable Eddington factors, $h_\nu$ is available, and hence $H_\nu(0) = h_\nu J_\nu(0)$ can be calculated directly. If Rybicki's method or the angle-dependent Feautrier equations are used, we can calculate $H_\nu(0) = \sum_m b_m\mu_m u(0, \mu_m, \nu)$. Alternatively, having $S_\nu(\tau_\nu)$ we can use the $\Phi$-operator [equation (2-61)] to find $F_\nu(0) = \Phi_0[S(\tau_\nu)]$; in practice this operation is done using a quadrature sum, for which several choices are available [see, e.g., (**141**; **246**; **8**, 33)]. If the flux is required at points internal to the atmosphere one may apply the operator $\Phi_\tau$ to $S_\nu$, or one may compute $v(\tau_{d\pm\frac{1}{2}}, \mu_m, \nu_n)$ from equation (6-15) and find $H_{d\pm\frac{1}{2},\,n} = \sum_m b_m\mu_m v_{d\pm\frac{1}{2},\,m,\,n}$ (note that this defines the flux at *midpoints* of the depth mesh).

# 7

# *Model Atmospheres*

## 7-1 The Classical Model-Atmospheres Problem: Assumptions and Restrictions

The *model-atmospheres problem* refers to the construction of mathematical models that provide a description of the physical structure of a stellar atmosphere and of its emergent spectrum. In its greatest generality, the problem is one of enormous complexity, and presents both physical and mathematical difficulties that are beyond solution at the present time. It is therefore necessary to make a number of simplifications, and to deal with idealized models that are rather high-order abstractions from reality. Such abstractions are useful inasmuch as they enhance our *insight* without overwhelming us with detail; yet it is important to state, at the outset, some of the restrictions we have imposed, not only because this helps to define the problem, but also as a reminder of the almost limitless numbers of fascinating research questions left to explore. The assumptions used in our work fall into several broad categories:

(a) *Geometry.* We assume that the atmosphere is composed of *homogeneous plane-parallel layers* when the thickness of the atmosphere is small



compared to the radius of the star, or (in 7-6) *homogeneous spherical shells* when the thickness is an appreciable fraction of the radius. The assumption of homogeneity makes the problem one-dimensional and thus *greatly* simplifies the analysis; but it excludes many interesting phenomena involved in small-scale structures seen in the solar atmosphere. For the stars we have almost no information about the homogeneity of the atmosphere [see, however, (**261**, Chap. 11)] and we can only hope that one-dimensional models yield some kind of "average" (in an ill-defined sense) information. However, because the "averaging" process is *nonlinear*, the question is really an open one, and it is not at all clear whether such models always do yield meaningful averages, (e.g., in chromospheres), although they *may* be satisfactory for some cases. In particular, in the solar atmosphere many of the inhomogeneities arise from hydrodynamic phenomena driven, ultimately, by the convection zone; for stars without strong convection zones, the atmospheres may indeed be homogeneous. (Counterexample: the Ap stars, which show gross variations of physical properties over their surface, presumably associated with the existence of strong magnetic fields).

(b) *Steady state.* We shall assume that the atmosphere is in a *steady state*, and shall avoid discussion of all time-dependent phenomena—e.g., stellar pulsations, shocks, transient expanding envelopes (novae, supernovae), heating by a binary companion, variable magnetic fields, etc. In this chapter we consider only *static* atmospheres; in Chapters 14 and 15 we extend the theory to steady flows (expanding atmospheres). We shall assume that the transfer equation is time-independent, and that level-populations are constant in time and are specified by statistical equilibrium equations (a special case being LTE) that equate the number of atoms leaving a level by all microprocesses to the number that return.

(c) *Momentum balance.* Having specified a steady state, we shall consider either *hydrostatic equilibrium* in which the static gas pressure distribution just balances gravitational forces, or one-dimensional, laminar, *steady flows.* Here we are ignoring the possibly large effects of magnetic forces: both large-scale (as in the Ap stars) and small-scale (e.g., in sunspots or in the concentrated knots of the general solar magnetic field). We further ignore the effects of small-scale motions such as waves, and larger scales such as super-granulation flows, convective cells, etc., as well as major tidal distortions in close binaries.

(d) *Energy Balance.* Usually we shall assume that the atmosphere is in *radiative equilibrium*, which again implies that it is *static*; in §7-3 we shall consider the effects of convection, but only in the roughest terms. In Chapter 15 we shall generalize to steady flow and include one-dimensional hydrodynamic work terms. The existence of complicated motions in the solar atmosphere is well documented observationally [see, e.g., (**694**, Chaps. 9 and 10) or (**244**, Chap. 5)] and, although data for stars are less complete,

there is little doubt that complex mass motions play an important role in the atmospheres of many stars (e.g., supergiants). But in its present state the theory is unequipped to handle with full consistency the details of energy exchange between the radiation field and hydrodynamic motions. Turbulent dissipation in convection; wave generation, propagation, and dissipation; effects of shear in rotating atmospheres; magnetic field effects; and a variety of other phenomena are all essentially overlooked! These are vital phenomena, for without them we cannot account for chromospheres and coronae (in this book we shall approach these regions from a semiempirical diagnostic view because we do not have an *ab initio* theoretical method). It remains true that *important* limits on our understanding of stellar atmospheres are imposed by our inability to handle the intricate interchange of energy between radiative and nonradiative modes, and that development of a satisfactory theory to handle such interactions is probably the most vital research frontier in this field of astrophysics.

It should be said, however (lest the reader receive an unduly gloomy picture of our efforts to date), that progress has been rapid, and continues at an accelerating rate, so that we may reasonably expect at least some of the inadequacies of the present-day theory to be ameliorated in the near future. Moreover, the framework imposed above does appear to yield many successful predictions of continuum features and line profiles for many (perhaps most) stars.

## 7-2 LTE Radiative-Equilibrium Models

In this section we develop the *methods* that can be used to construct planar, static, radiative-equilibrium models assuming LTE; the *results* of such calculations will be described in §7-4. As was discussed in Chapter 5, the assumption of LTE vastly simplifies the computation (as one can see by comparing the methods of this section with those of §7-5). We criticized the use of LTE because it does not give an accurate description of the interactions of radiation and matter in stellar atmospheres, and is totally deficient in many important conceptual points (especially regarding line-formation). But on the pragmatic side, LTE models allows treatment of many effects (e.g., line-blanketing) that are of importance in the application of the results of stellar atmospheres computations to the interpretation of photometric indices, stellar temperatures and luminosities, etc., but that still lie beyond the present capabilities of a non-LTE calculation. In a sense, then, the two approaches are complementary: the non-LTE theory provides deep physical insight while LTE allows a preliminary assessment of complexities in the models. Of course the end goal will be to have non-LTE models that are as "refined" as any LTE model can be.



### THE OPACITY AND EMISSIVITY: CONTINUA AND LINE-BLANKETING

The frequency variation of the opacity and emissivity in stellar atmospheres plays a key role in determining the nature of the emergent spectrum. For example, the sharp decrease in flux shortward of about $\lambda 3650$ Å in A-stars can be explained by the large jump in the opacity caused by photoionizations from the $n = 2$ state of hydrogen. Because the material becomes more opaque, we see less deeply into the atmosphere, and therefore receive energy only from the outer, cooler, layers. We have already seen (Chapter 3), that we *cannot* reduce the problem of an atmosphere with a nongrey opacity to the grey problem by any choice of average opacity, and we must, therefore, make allowance for the detailed frequency-dependence of the absorption coefficient from the outset. At the very minimum we must treat the opacity variation in the *continuum*, which accounts for the gross features of the energy distribution in the emergent spectrum; in more refined work we must also include the effects of *lines*.

The opacity at any given frequency contains contributions from all possible transitions (bound–bound, bound–free, free–free) of all chemical species that can absorb photons at that frequency. From equations (5-53) and (5-60) we see that the direct absorption coefficient for process $(i \to j)$ from level $i$ is $n_i\alpha_{ij}(\nu)$. Stimulated emissions return energy to the beam at a rate proportional to $I_\nu$; hence (assuming identity of the emission and absorption profiles) we correct the opacity by subtracting stimulated emissions from the absorptivity. In view of equations (5-54) and (5-64), the correction is $n_j\alpha_{ij}(\nu)G(\nu)$ where $G(\nu) = g_i/g_j$ or $G(\nu) = (n_i/n_j)^* \exp(-h\nu/kT)$ for bound–bound or bound–free processes respectively. Let $n_i^*$ denote the LTE population of state $i$ computed from the usual Saha–Boltzmann formulae [equation (5-14)] using the *actual* ion density. Then summing over all levels and processes we have the *non-LTE opacity*

$$\chi_\nu = \sum_i \sum_{j>i} [n_i - (g_i/g_j)n_j]\alpha_{ij}(\nu) + \sum_i (n_i - n_i^* e^{-h\nu/kT})\alpha_{i\kappa}(\nu) + \sum_\kappa n_e n_\kappa \alpha_{\kappa\kappa}(\nu, T)(1 - e^{-h\nu/kT}) + n_e\sigma_e \tag{7-1}$$

where the four terms represent, respectively, the contributions of bound–bound, bound–free, and free–free absorptions, and of electron scattering (other scattering terms—e.g., Rayleigh scattering—may also be added). To calculate the *spontaneous thermal emission* (non-LTE) we use the rates derived in equations (5-55) and (5-62) to write

$$\eta_\nu = (2h\nu^3/c^2)\left[\sum_i \sum_{j>i} n_j(g_i/g_j)\alpha_{ij}(\nu) + \sum_i n_i^*\alpha_{i\kappa}(\nu)e^{-h\nu/kT} + \sum_\kappa n_e n_\kappa \alpha_{\kappa\kappa}(\nu, T)e^{-h\nu/kT}\right] \tag{7-2}$$

The three terms again describe bound–bound, bound–free, and free–free processes. Emission from continuum scattering terms will be written separately in the transfer equations. Equations (7-1) and (7-2) apply in the non-LTE case; *if we assume LTE* they simplify to

$$\chi_\nu^* = \left[\sum_i \sum_{j>i} n_i^* \alpha_{ij}(\nu) + \sum_i n_i^* \alpha_{i\kappa}(\nu) + \sum_\kappa n_e n_\kappa \alpha_{\kappa\kappa}(\nu, T)\right] \times (1 - e^{-h\nu/kT}) + n_e \sigma_e \tag{7-3}$$

and

$$\eta_\nu^* = (2h\nu^3/c^2) e^{-h\nu/kT} \times \left[\sum_i \sum_{j>i} n_i^* \alpha_{ij}(\nu) + \sum_i n_i^* \alpha_{i\kappa}(\nu) + \sum_\kappa n_e n_\kappa \alpha_{\kappa\kappa}(\nu, T)\right] \tag{7-4}$$

Writing $\kappa_\nu^* = \chi_\nu^* - n_e\sigma_e$, we see that $\eta_\nu^* = \kappa_\nu^* B_\nu$, as expected from the Kirchhoff–Planck relation [equation (2-6)].

In LTE, the occupation numbers $n_i^* = n_i^*(N, T)$, hence $\chi_\nu^* = \chi_\nu^*(N, T)$ and $\eta_\nu^* = \eta_\nu^*(N, T)$, which simplifies the computation and allows great numbers of absorption and emission processes to be included easily. The basic free parameters that enter the calculation are those describing the *chemical composition* of the material. In different spectral types, different absorbers will dominate, depending upon the ionization and excitation state of the material. Thus for stars of solar temperature and cooler, the dominant bound–free absorption process is from the $H^-$ ion; for A-stars it is from neutral H; in the B-stars He I begins to make significant contributions; in the O-stars He II and numerous light-element ions (e.g., of C, N, O, Ne, Si) play an important role (**101**), (**319**). In the later-type stars a wide variety of negative ions of atoms and molecules contribute [(**109**), (**73**, Chap. 4), (**644**)]. In general literally hundreds or thousands of levels may contribute; only in LTE can this much detail be handled, and even then extensive computations are required [see, e.g., (**504**)]. Free–free absorption from $He^+$, He and H is important in the O-stars; for the A-stars the main free–free contribution is from H; in the sun it is from $H^-$; and in the M-stars $H_2^-$ free–free absorption becomes significant. Electron scattering is a major opacity source in the O-stars, while Rayleigh scattering by H and $H_2$ contributes to the opacity in the atmospheres of intermediate-temperature stars (spectral types G and K). Very elaborate and comprehensive opacity calculations have been made by workers of the groups at Kiel and Los Alamos, who have published extensive graphs and tables of results [the reader should examine these carefully; see (**651**; **97**; **638**, 181–199; **184**)]. The bulk of the Los Alamos work pertains to stellar interiors, but some of the results are relevant to stellar atmospheres; a very complete discussion of the techniques used in these computations may be found in (**14**, Chap. 3). For continua the calculations may be laborious, but are straightforward in principle.



In addition to continua, the opacity of stellar material contains contributions from *thousands to millions* of spectral lines, both atomic and molecular. Bound–bound opacities are significant for stars of all spectral types. For the earliest spectral types the resonance lines of H, He I, He II, and ions of light elements dominate in the ultraviolet. For A-stars the hydrogen Lyman and Balmer lines are important. For solar-type stars the important effects come from lines of neutral and singly ionized metals and other atoms of moderate atomic weight. In later types, molecular bands (CN, CO, $H_2O$, etc.) dominate. Again, the chemical composition of the gas is a fundamental parameter in setting the line opacity. In addition, parameters that determine linewidths—e.g., macroscopic velocities in the atmosphere (the so-called "microturbulence"; see §10-3)—also enter.

The effects of bound–bound absorptions upon a stellar atmosphere are referred to as *line-blanketing*, and play a crucial role in determining both the emergent energy distribution and the physical structure of the atmosphere. The temperature gradient in the atmosphere implies that, in the opaque lines, the layers from which we receive radiation will be cooler, and hence radiate less energy. The presence of numerous dark spectral lines within a given photometric band (established, e.g., by a filter) obviously directly affects the measured flux. This effect is called the *blocking effect*. Because the total flux in the atmosphere must be conserved, the flux blocked by the lines must emerge at other frequencies, and the energy emitted in the continuum bands into which it is redistributed rises above the value it would have had in the absence of lines. Furthermore, because the bandwidth of the spectrum in which energy transport occurs readily is restricted by lines, steeper temperature gradients are necessary to drive the flux through; as a result, temperatures in underlying layers rise, leading to the *backwarming effect*. Finally, the *lines alter the temperature in the outermost layers* of the atmosphere. We shall study these effects further in §7-4; it is clear from what has been said that it is important to include bound–bound opacities in the calculation, and we ask here how this may be done.

The most straightforward method of treating lines is the *direct approach* in which one includes enough frequency points in the calculation to describe the profiles of the lines under consideration. The full frequency and depth-variation of the absorption coefficient is taken into account by this method, which may be applied when the spectrum is dominated by just a few lines. The direct method suffers from the disadvantage that for many stars the line spectrum (e.g., in molecular bands that contain millions of lines) is so complex that a detailed description is prohibitively expensive in computing time, and we must therefore consider alternatives, which we may categorize as *statistical methods*. Here one attempts to replace the complicated frequency variation of the line opacity in a given band (see Figure 7-1) by a small number of parameters. The simplest possible description is to reduce

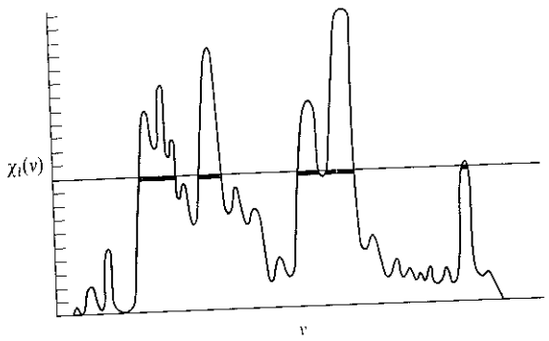



FIGURE 7-1
Schematic absorption coefficient of overlapping spectrum lines. A very large number of frequency points would be required to describe the detailed variation of the opacity.

all the information to a *single number*, given by a *mean opacity*; in particular we could consider either the Planck mean [cf. equation (3-28)] or the Rosseland mean [cf. equation (3-26)]. As might be expected, this approach is not really satisfactory (for precisely the same reasons it fails for the continuum). For example, while it is quite adequate to use the Rosseland mean at depth where the diffusion approximation holds (as is done to include line absorption in stellar interiors opacities), this method tends to underestimate the opacity near the surface and yields poor approximations to the actual energy balance there (**126**). The Planck mean fails to yield the diffusion limit for the flux at depth, and grossly overestimates the opacity at the surface; this leads to substantial errors in predicted fluxes and temperature structure of the model.

Recognizing the inadequacy of a single mean opacity we instead replace the detailed spectrum by a smooth *opacity distribution function* (which is a generalization of the classical picket-fence model described in §7-4), as in Figure 7-2. We consider a narrow enough band to assure that the exact position of the line is not important (i.e., other properties such as the continuum opacity or Planck function do not vary much over the band). We then could, for example, find the fraction of the band covered by lines with opacity $\chi_l(\nu) \geq$ some chosen value $X_i$, and plot a graph of this fraction against $X_i$. The result is a smooth curve that can be well approximated by a small number of subintervals (possibly of differing widths) containing constant opacities appropriate to the curve. This procedure may be carried out for a mesh of temperatures and densities to produce a description of the variation of the line opacity through the atmosphere. A critical study of this approach (**126**) shows that opacity distribution functions yield excellent results, and reproduce both the emergent fluxes and physical atmospheric structure given by detailed direct calculations to satisfactory accuracy.



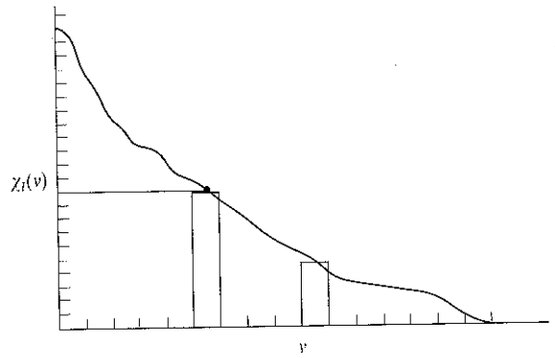



FIGURE 7-2
Schematic opacity distribution function of the spectrum in Figure 7-1. A relatively small number of representative opacities suffice to describe this smooth distribution.

The main limitation of the opacity distribution function approach is that it implicitly assumes that the positions of the lines (in frequency) do not change markedly as a function of depth, measured in units of a photon mean-free-path (i.e., unit optical depth in the continuum). It is crucial to the transfer problem whether a line in one layer of the atmosphere coincides in frequency with a line or with a continuum band in an overlying layer, for photons might freely escape in the latter case, but not the former. Marked variations in the line spectrum, which invalidate the opacity distribution function approach, can occur in a number of situations—for example, the following. (a) Molecular bands of two species may overlap; one species may show a rapid decrease or increase with depth relative to the other. Even though the total opacity of the two bands together might not change, the positions of the two sets of lines could be radically different. (b) A strong shock in the atmosphere might produce an abrupt change in the excitation–ionization state of the gas over a small distance. The line spectra through the shock front might change radically. (c) Velocity shifts in expanding atmospheres systematically move lines away from their rest positions; this strongly affects momentum and energy balance in the material (cf. §§14-1 and 15-4). In such cases one must employ either the direct approach, or a generalization of the statistical approach that in some way allows for the changes in the frequency positions of the lines. An alternative approach, called the *opacity sampling* technique (based on a random-sampling procedure) has recently been suggested (**585**); although this method appears computationally more costly than the opacity distribution function method, it also appears that it may not suffer from the limitations just described, and should be tested further.

HYDROSTATIC EQUILIBRIUM

In a static atmosphere, the weight of the overlying layers is supported by the total pressure, and it is this balance, in essence, that determines the density structure of the medium. Thus

$$\nabla p = \rho \mathbf{g} \tag{7-5}$$

where the total pressure $p = p_g + p_R$ (dynes cm$^{-2}$); the gas pressure $p_g \equiv NkT$; the radiation pressure $p_R \equiv (4\pi/c) \int K_\nu \, d\nu$; $\mathbf{g}$ is the surface gravity (regarded as a fundamental parameter describing the atmosphere). Here $\rho$ is the mass density (gm cm$^{-3}$) which, using the notation of 5-2, can be written

$$\rho = (N - n_e) m_{\mathrm{H}} \sum_k \alpha_k A_k \equiv (N - n_e)\bar{m} \tag{7-6}$$

where $m_{\mathrm{H}}$ is the mass of a hydrogen atom, and $A_k$ is the atomic weight of chemical species $k$ with fractional abundance $\alpha_k$. If we define the *column mass* $m$ (gm cm$^{-2}$) measured from the outer surface inward as our new independent variable—i.e.,

$$dm = -\rho \, dz \tag{7-7}$$

then we may rewrite equation (7-5) as $dp/dm = g$, which yields an *exact integral* $p(m) = gm + c$. It is obviously advantageous to be able to write such a result, so we shall use $m$ as the independent variable henceforth; the choice of $m$ instead of $z$ has no significant effect on the transfer equation. Using equation (2-77b) for the radiation pressure gradient we can rewrite equation (7-5) in another useful form:

$$(dp_g/dm) = g - (4\pi/c) \int_0^\infty (\chi_\nu/\rho) H_\nu \, d\nu \tag{7-8}$$

which shows that radiation forces tend to cancel gravitational forces, and lead to a smaller pressure gradient in the atmosphere. Put another way, the material tends to "float" upon the radiation field. As was shown in Chapter 1, the radiation force is related to the flux through the atmosphere, and we can thus see that for a given $T_{\mathrm{eff}}$ there will be some lower bound on $g$ below which radiation forces exceed gravity and blow the material away. Specifically, Underhill (**633**) showed that gravity forces will exceed radiation forces only if $g \gtrsim 65 \, (T_{\mathrm{eff}}/10^4)^4$ cm sec$^{-2}$. Clearly radiation pressure forces are negligible for the sun ($T_{\mathrm{eff}} \approx 6 \times 10^3$, $g \approx 3 \times 10^4$) but become very important for an O-star ($T_{\mathrm{eff}} \approx 4 \times 10^4$, $g \approx 10^4$) and in supergiants where $g$ is quite low (and indeed approaches $g_{\mathrm{crit}}$). In fact, as we shall show in Chapter 15, for some O-stars the radiation forces on spectrum lines in stellar winds *exceed* $g$ and accelerate the material to very large velocities ($\sim 3000$ km sec$^{-1}$).



*Exercise 7-1*: Consider *fully ionized* stellar material of hydrogen and helium (abundance $Y$). (a) Show that $n_e\sigma_e/\rho$, which provides a lower bound on the opacity, is $\sigma_e(1 + 2Y)/m_{\mathrm{H}}(1 + 4Y)$. (b) Take advantage of the grey nature of electron scattering to show that, if gravity is to exceed radiation forces, then $g$ must be $\geqslant g_{\mathrm{crit}}$ where $g_{\mathrm{crit}} = \sigma_e(1 + 2Y)(\sigma_R T_{\mathrm{eff}}^4)/[cm_{\mathrm{H}}(1 + 4Y)]$. (c) Re-express this result to show that the luminosity $L$ of the star must be $\leqslant L_{\mathrm{crit}} \approx 3.8 \times 10^4 \, (\mathscr{M}/\mathscr{M}_\odot) L_\odot$.

For computational purposes we can rewrite (7-5) as a difference equation connecting the depths specified by column masses $m_d$ and $m_{d+1}$, namely

$$N_d kT_d - N_{d-1} kT_{d-1} + (4\pi/c) \sum_{n=1}^{N} w_n (f_{dn} J_{dn} - f_{d-1,n} J_{d-1,n}) = g(m_d - m_{d-1}) \tag{7-9}$$

Here $K_\nu$ is expressed in terms of the mean intensity and a variable Eddington factor; i.e., $K_\nu = f_\nu J_\nu$. We can obtain a starting value from equation (7-8) by assuming that the radiation force remains constant from the boundary surface upward, and thus

$$N_1 kT_1 = m_1 \left[ g - (4\pi/c) \sum_n w_n (\chi_{1,n}/\rho_1) h_n J_{1,n} \right] \tag{7-10}$$

Equations (7-5) through (7-10) are valid for both LTE and non-LTE atmospheres.

Notice that, if we knew the temperature structure $T(m)$, and could either (a) ignore radiation forces or (b) estimate them, using equation (7-8), as $(\bar{\chi}/\rho)(\sigma_R T_{\mathrm{eff}}^4/c)$ where $\bar{\chi}$ is a suitable mean opacity, then we immediately could derive the density structure $N(m)$. From this we could calculate $\chi_\nu^*(N, T)$, $\eta_\nu^*(N, T)$, solve the transfer equation, and thus determine all model properties of interest. Of course in general we do *not* know the temperature structure, and we must now address the issue of how it is to be determined.

RADIATIVE EQUILIBRIUM:
TEMPERATURE-CORRECTION PROCEDURES

For a given temperature distribution, the equation of hydrostatic equilibrium can be integrated as described above, and opacities and emissivities

derived. The radiation field then follows from a solution of the LTE transfer equation

$$\partial^2(f_\nu J_\nu)/\partial\tau_\nu^2 = J_\nu - (\eta_\nu^* + n_e\sigma_e J_\nu)/\chi_\nu^* = (1 - n_e\sigma_e/\chi_\nu^*)J_\nu - (\kappa_\nu^*/\chi_\nu^*)B_\nu \tag{7-11}$$

at all frequencies and depths using the techniques described in Chapter 6. For an atmosphere in radiative equilibrium the total energy absorbed by the material must equal that emitted, hence in LTE

$$4\pi\int_0^\infty [\eta_\nu^* - (\chi_\nu^* - n_e\sigma_e)J_\nu]\,d\nu = 4\pi\int_0^\infty \kappa_\nu^*(B_\nu - J_\nu)\,d\nu = 0 \tag{7-12}$$

or, in discrete form, (and allowing for departures from LTE),

$$4\pi\sum_n w_n[\eta_n - (\chi_n - n_e\sigma_e)J_n] = 0 \tag{7-13}$$

In radiative equilibrium the total flux $4\pi H = \sigma_R T_{\text{eff}}^4 =$ constant, and we may choose it (or $T_{\text{eff}}$) as another fundamental parameter characterizing the model. Now in general we do not know *the* temperature distribution that produces radiative equilibrium, and using our present *estimate* of $T(m)$ we will normally find that the radiation field does *not* satisfy equation (7-12) or (7-13). It is therefore necessary to adjust $T(m)$ iteratively in such a way that the radiation field does ultimately satisfy the requirement of energy balance. The determination of $T(m)$ is, in fact, the very heart of the problem of constructing LTE models. There are basically two strategies we may use: (a) *temperature correction procedures*, and (b) solution of the transfer equation *subject to a constraint of radiative equilibrium*. In temperature correction procedures one attempts to use information about the radiation field calculated from a given $T(m)$ in an *a posteriori* fashion to estimate a change $\Delta T(m)$ that will cancel out the errors found in the flux and in the flux derivative [equivalent to equations (7-12) and (7-13); see equation (2-71)]. In the second approach, one attempts from the outset to formulate the transfer equation in such a way that the resulting radiation field will *automatically* satisfy radiative equilibrium. The first approach (corrections) was historically the one originally used to solve the nongrey atmospheres problem, and the methods are often quite ingeniously constructed. The second approach (constraints) is more subtle and powerful, and overcomes inadequacies fatal to "correction" procedures in the non-LTE case, thus allowing a deep penetration into problems of considerable complexity. Ironically, the roots of the idea of using "constraints" are to be found in the methods used to solve the *grey* problem. We first consider temperature correction procedures.

The first, and most obvious, method is the so-called *lambda-iteration procedure*. Here we suppose that from a given run of $T_0(m)$ we have, in effect, computed $J_\nu = \Lambda_{\tau_\nu}[B_\nu(T_0)]$, hence the name of the method, and that equation (7-12) is *not* satisfied. We then assume that the run of $T(m)$ that *does* satisfy the condition of radiative equilibrium is $T(m) = T_0(m) + \Delta T(m)$, and require that



$$\int_0^\infty \kappa_\nu^* B_\nu(T_0 + \Delta T)\,d\nu = \int_0^\infty \kappa_\nu^* J_\nu\,d\nu \tag{7-14}$$

Expanding $B_\nu(T_0 + \Delta T) \approx B_\nu(T_0) + \Delta T(\partial B_\nu/\partial T)$ we find

$$\Delta T \approx \int_0^\infty \kappa_\nu^*[J_\nu - B_\nu(T_0)]\,d\nu \Big/ \int_0^\infty \kappa_\nu^*(\partial B_\nu/\partial T)_{T_0}\,d\nu \tag{7-15}$$

It must be emphasized that $J_\nu$ in equation (7-15) denotes the value *already computed* from $B_\nu(T_0)$. If one carries through the process and recomputes a new model with the new temperature distribution, some improvement in satisfying equation (7-12) usually will be found. However, the procedure suffers from several severe defects.

(a) Because $J_\nu = \Lambda_{\tau_\nu}[B_\nu(T_0)] = B_\nu(T_0) + 0(e^{-\tau_\nu})$, it is clear that at depth the temperature correction goes rapidly to zero, no matter how bad the solution actually is at those points. We found a similar result in the grey problem.

(b) If the frequency variation of $\kappa_\nu^*$ is such that the opacity is much larger (say several orders of magnitude) at some frequencies than at others, the method again fails. The reason is that in the opaque frequency bands the contribution to the numerator vanishes as $\tau_\nu \to 1$ while the contribution to the denominator swamps that of all other bands. In effect the Λ-iteration procedure is effective only over $\Delta\tau_\nu \sim 1$ for the most opaque frequencies.

(c) Equation (7-12) places a condition only on the flux *derivatives*; hence we have no way of specifying the actual value of the flux to which the solution converges (if it does).

(d) The real failure of the Λ-iteration procedure is that it *ignores* the effect that $\Delta T$, computed at some depth $\tau$, has on $J_\nu(\tau')$ at all other depths (i.e., $J_\nu$ is presumed to be *fixed*). This oversight necessarily leads to spurious values of $\Delta T$. Actually $J_\nu(\tau_\nu') = \Lambda_{\tau_\nu'}[B_\nu(T + \Delta T)]$, which means we should, in reality, solve an *integral equation for* $\Delta T$; we shall return to this point later.

When the reasons for the failure of the Λ-iteration procedure were understood, it was realized that methods were needed that made use of information about errors in the flux itself (which gives direct information about the temperature gradient at depth) as well as in the flux derivative. One method of doing this was suggested by Lucy (**283**, 93), who generalized the method devised by Unsöld for the grey problem (see §3-3) to the nongrey case. If we use a Planck mean [equation (3-23)] optical depth scale $d\tau \equiv -\kappa_P^*\,dz$, then exact frequency integrals of the moment equations (using quantities

without subscripts to denote frequency-integrated variables) are

$$(dH/d\tau) = (\kappa_J^*/\kappa_P^*)J - B \tag{7-16}$$

and

$$(dK/d\tau) = (\chi_F^*/\kappa_P^*)H \tag{7-17}$$

where $\kappa_J^*$ is the absorption mean [equation (3-32)] and $\chi_F^*$ is the flux mean [equation (3-21)]; note that the scattering coefficient is included in $\chi_F^*$ but not in the other means. Then relating $K$ to $J$ with the Eddington approximation, equations (7-16) and (7-17) are combined to give an expression for $B(\tau)$, which, treated as a perturbation equation for a correction $\Delta B(T) = 4\sigma_R T^3 \Delta T/\pi$ gives finally

$$\Delta B(\tau) = -d(\Delta H)/d\tau + (\kappa_J^*/\kappa_P^*)\left[3\int_0^\tau (\chi_F^*/\kappa_P^*)\,\Delta H(\tau')\,d\tau' + 2\,\Delta H(0)\right] \tag{7-18}$$

Here $\Delta H(\tau) \equiv H - H(\tau)$. The first term on the righthand side of equation (7-18) is the correction predicted by the Λ-iteration procedure; the other terms introduce new information that gives nonnegligible values of $\Delta B$ at depth and produces a response to flux errors at the surface. Experience has shown the Unsöld–Lucy procedure to be quite effective in constructing LTE radiative-equilibrium models (but it has no obvious generalization to the non-LTE case).

*Exercise 7-2*: Derive equations (7-16) and (7-17) and, applying reasoning similar to that yielding equation (3-44), derive equation (7-18).

Another very clever and quite useful method of calculating temperature corrections was suggested by Avrett and Krook (**55**), who introduce perturbations to both the temperature and the optical depth scale. That is, we suppose that the current temperature distribution $T_0(t)$ is related to the desired temperature distribution $T(\tau)$ (which yields radiative equilibrium) by a pair of relations: $T = T_0 + T_1$ and $\tau = t + \tau_1$. The transfer equation is then expanded to first order in the perturbations $\tau_1$ and $T_1$, and by taking moments of the resulting first-order equation of the perturbation expansion, equations are derived for $\tau_1$ and $T_1$. These equations [extended to allow for scattering terms (**421**) and with an improved closure relation (**351**)] are

$$\tau_1' + \tau_1\left(\int_0^\infty \chi_\nu H_\nu^0\,d\nu \Big/ \int_0^\infty \chi_\nu H_\nu^0\,d\nu\right) = (1 - \mathscr{H}/H^0) - 3^{-\frac{1}{2}}\int_0^\infty \chi_\nu[J_\nu^0 - B_\nu(T_0)]\,d\nu \Big/ \int_0^\infty \chi_\nu H_\nu^0\,d\nu \tag{7-19}$$



and

$$T_1 = \Big\{(1+\tau_1)\int_0^\infty \chi_\nu(1-\rho_\nu)[J_\nu^0 - B_\nu(T_0)]\,d\nu - 3^{\frac{1}{2}}(1-\mathscr{H}/H^0)\int_0^\infty \chi_\nu(1-\rho_\nu)H_\nu^0\,d\nu + \tau_1'\int_0^\infty [\chi_\nu'(1-\rho_\nu) - \chi_\nu\rho_\nu'][J_\nu^0 - B_\nu(T_0)]\,d\nu\Big\} \Big/ \int_0^\infty \chi_\nu(1-\rho_\nu)(\partial B_\nu/\partial T)_0\,d\nu \tag{7-20}$$

where the prime denotes the derivative $d/dt$; quantities with superscript or subscript zero denote current values; $H^0 \equiv \int_0^\infty H_\nu^0\,d\nu$; $\mathscr{H}$ denotes the nominal flux $(\sigma T_{\text{eff}}^4/4\pi)$; and $\rho_\nu \equiv \sigma_\nu/\chi_\nu = \sigma_\nu/(\kappa_\nu + \sigma_\nu)$. Equation (7-19) is a linear first-order differential equation for $\tau_1$ that may be integrated straightaway starting with $\tau_1(0) = 0$; given values for $\tau_1$ and $\tau_1'$, one then computes $T_1$ from equation (7-20) and hence obtains $T(\tau) = T_0(t) + T_1(t)$ at $\tau = t + \tau_1(t)$. Experience shows that the $\tau_1$ correction leads to revisions of the temperature scale at depth, while $T_1$ is most important at the surface.

Both the Unsöld–Lucy and the Avrett–Krook procedures have been widely applied and have been proven quite successful in the construction of LTE radiative-equilibrium models. With care in calculation, these methods produce models with errors in $|\Delta F/F|$ and $|d \ln F/d\tau|$ of the order of 0.1 percent. But despite the fact that temperature correction procedures are easy to use and work fairly well, they have a number of serious drawbacks that render them ineffective against non-LTE problems. We shall describe these briefly here to motivate the discussion of the constraint-type procedures which, although first developed to attack the non-LTE case, also are extremely effective for LTE models, and are now the preferred methods.

First, temperature-correction methods tend to stabilize rather than converge if there are large variations in the frequency dependence of the opacity. This failure is particularly troublesome when we attempt to construct models including spectral lines or a major continuum jump (e.g., the hydrogen Lyman continuum). In such cases temperature correction procedures leave the temperature structure of the outermost layers essentially undetermined, for the energy balance there is established entirely by these opaque transitions (the optically thin regions already have radiation fields that are *fixed*). Second, these methods tacitly assume that the *temperature* is the crucial variable, and cannot deal effectively with cases where the radiation field is only weakly coupled to the local thermal pool—e.g., in atmospheres where scattering terms dominate or where non-LTE line formation occurs (we shall see why this is so in §7-5 and Chapters 11 and 12). Finally, these methods are not sufficiently accurate. Although errors of a tenth of a percent seem

small, it must be realized that there is a close similarity in the requirements of radiative equilibrium and statistical equilibrium, and that errors of this size may be totally unacceptable in the latter context. Specifically, suppose we consider a ground-state continuum that dominates the opacity by orders of magnitude. Then energy balance requires $\int_{\nu_0}^{\infty} \alpha_\nu{}^1 J_\nu \, d\nu = \int_{\nu_0}^{\infty} \alpha_\nu{}^1 B_\nu \, d\nu$ while the net radiative rate in the statistical equilibrium calculation is of the form $\int_{\nu_0}^{\infty} \alpha_\nu{}^1 [b_1 J_\nu - B_\nu] \nu^{-1} \, d\nu$ (where $b_1$ denotes $n_1/n_1^*$ of the ground state). If $b_1 \approx 1$, these two criteria differ only by the (weak!) "profile" function $\nu^{-1}$. In the limit that $h\nu_0/kT \gg 1$ and both $B_\nu$ and $J_\nu$ have a characteristic frequency variation of $\exp(-h\nu/kT)$, both pairs of integrals are strongly dominated by contributions from $\nu \approx \nu_0$, and hence one of them becomes essentially redundant. Thus an error of a few tenths of a percent in energy balance implies a similar error in the net radiative rate; because the radiative rates may exceed the collisional rates by orders of magnitude (recall the discussion of §7-3), such errors may overwhelm all other terms in the rate equation, and lead to false equilibration. We now turn to a discussion of methods that treat the condition of radiative equilibrium as a *constraint*; these overcome all of the shortcomings described above.

The essence of the constraint approach is *to build the requirements of radiative equilibrium directly into the transfer equation* in such a way as to solve both problems *simultaneously*. These methods deal directly with the global nature of the radiation field; i.e., they account for the effect that a temperature change at one point in the atmosphere produces upon the radiation field at all other points, and vice versa. To simplify the discussion in the remainder of this subsection we shall ignore scattering terms in the source function. One such procedure was proposed by E. Böhm-Vitense (**283**, 99) who suggested that the integral equation for $\Delta T$,

$$\int_0^\infty \kappa_\nu^* B_\nu(T + \Delta T) \, d\nu = \int_0^\infty \kappa_\nu^* \Lambda_{\tau_\nu}[B_\nu(T + \Delta T)] \, d\nu \tag{7-21}$$

be solved directly; we shall recast her discussion into slightly different terms using current notation (see also **32**). To solve an integral equation of this form we first construct a matrix representation for the $\Lambda$-operator; we introduce a discrete set of points $\{\tau_{\nu,d}\}$, $(d = 1, \ldots, D)$, at which we wish to determine the solution, and we represent the variation of $B_\nu(\tau_\nu)$ *analytically* in terms of interpolation (*basis*) functions on this mesh. The integral

$$J_\nu(\tau_\nu) = \frac{1}{2} \int_0^\infty E_1|t_\nu - \tau_\nu| \, B_\nu(t_\nu) \, dt_\nu \tag{7-22}$$

is then calculated *analytically* for these basis functions to produce the system

$$J_{d,\nu} = \sum_{d'=1}^{D} \Lambda_{dd',\nu} B_{d',\nu} + M_{d,\nu}, \quad (d = 1, \ldots, D) \tag{7-23}$$



where the $M_{d,\nu}$ represent the contribution of the interval $(\tau_{D,\nu}, \infty)$ to the integral. Substituting equation (7-23) into (7-21), and (a) assuming $\kappa_\nu^*$ is unchanged by $\Delta T$, (b) writing $B_\nu(T + \Delta T) \approx B_\nu(\tau) + (\partial B_\nu/\partial T)\,\Delta T$, and (c) introducing a frequency quadrature $\{\nu_n\}$, $(n = 1, \ldots, N)$, we find a set of linear equations for the values of $\Delta T_d$:

$$\sum_{d'=1}^{D} \left[ \sum_{n=1}^{N} w_n \kappa_{d'n}^* (\partial B/\partial T)_{d'n} (\delta_{dd'} - \Lambda_{dd'n}) \right] \Delta T_{d'} = \sum_n w_n \kappa_{dn}^* \left[ M_{dn} + \sum_{d'=1}^{D} (\Lambda_{dd'n} - \delta_{dd'}) B_{d'n} \right], \quad (d = 1, \ldots, D) \tag{7-24}$$

*Exercise 7-3*: Verify equation (7-24).

The solution of this system yields the change in the temperature consistent with the global properties of the radiation field. Because our expansion of $B_\nu$ is only linear, the system has to be iterated to convergence by using the new temperatures to recalculate $B_\nu$, $(\partial B_\nu/\partial T)$, $\kappa_\nu^*$, etc., and re-solving the system; if assumption (a) is valid, we would expect quadratic convergence. There are some defects to this approach. (1) The computation of the $\Lambda$ matrix is cumbersome and costly, and (because $\kappa_\nu^*$ really *is* a function of $T$) must be done again for each iteration. (2) It is possible to calculate the response of the $\Lambda$ matrix to changes in $\tau_\nu$ (arising from changes in $\kappa_\nu^*$ caused by the temperature change) but this again is extremely cumbersome and costly (also, there are problems of stability) [see, e.g., (**347**; **575**)]. The method described in the final subsection of this section overcomes these difficulties. (3) As originally formulated, and as described thus far, the method does not force convergence to a prespecified flux. This may be done by applying the diffusion approximation at the lower boundary and demanding the correct flux transport (**32**). Thus, we write for $\tau_\nu > \tau_{D,\nu}$

$$B_\nu(\tau_\nu) = B_\nu(T_D) + \frac{1}{\kappa_\nu^*} \frac{\partial B_\nu}{\partial T} \left| \frac{dT}{dz} \right| (\tau_\nu - \tau_{D,\nu}) \tag{7-25}$$

$$I(\tau_\nu, \mu) \approx B_\nu(T_D) + \frac{\mu}{\kappa_\nu^*} \frac{\partial B_\nu}{\partial T} \left| \frac{dT}{dz} \right| \tag{7-26}$$

Integrating against $\mu$, and over all frequencies, we find

$$H = \frac{1}{3} \left( \int_0^\infty \frac{1}{\kappa_\nu^*} \frac{\partial B_\nu}{\partial T} \, d\nu \right) \left| \frac{dT}{dz} \right| \tag{7-27}$$

which fixes $dT/dz$ in equation (7-25), and introduces the flux into the quantities $M_{d,\nu}$; note the similarity of this device to that used in the grey problem.

*Exercise 7-4*: (a) Evaluate the elements $\Lambda_{dd'}$ by assuming a piecewise linear interpolation for $B_\nu(\tau_\nu)$ on a discrete grid; i.e., on

$$[\tau_d, \tau_{d+1}], \text{ set } B(\tau) = [B_d(\tau_{d+1} - \tau) + B_{d+1}(\tau - \tau_d)]/(\tau_{d+1} - \tau_d)$$

(b) Evaluate $M_d$ using equations (7-25) and (7-27).

*Exercise 7-5*: Using a digital computer, program the solution of equation (7-24) for an opacity step function: $\kappa_\nu = \kappa_0$ for $(\nu \leqslant \nu_0)$, $\kappa_\nu = \alpha\kappa_0$ for $\nu > \nu_0$; parameterize the problem in terms of the value of $\beta = h\nu_0/kT_{\text{eff}}$, which specifies the frequency of the step. Notice that the $\Lambda$ matrix and $M$ vector are independent of frequency (though different) in the two ranges $(\nu \leqslant \nu_0)$, $(\nu > \nu_0)$; hence the integrals over frequency of $B_\nu$ and $\partial B_\nu/\partial T$ can be done analytically and expressed in terms of elementary functions [by using the known result for the complete interval $(0, \infty)$ and appropriate expansions for $(0, \nu_0)$ when $\beta \ll 1$ and for $(\nu_0, \infty)$ when $\beta \gg 1$] or in terms of Debye integrals (**4**, 998). Solve the problem for several values of $\alpha$ and $\beta$, starting from the grey temperature distribution (on the Rosseland mean scale); compare your results with those in (**128**; **605**; **38**).

A second constraint procedure was suggested by Feautrier (**283**, 108; **210**). Noting that radiative equilibrium implies that

$$\sum_n w_n \kappa^*_{dn} J_{dn} \Big/ \sum_n w_n \kappa^*_{dn} B_{dn} \equiv 1 \tag{7-28}$$

at all depths $d$, he solved the transfer problem [equation (6-30) or (6-42)] with the source function

$$S_{dn} = B_{dn} \cdot \left( \sum_{n'} w_{n'} \kappa^*_{dn'} J_{dn'} \Big/ \sum_{n'} w_{n'} \kappa^*_{dn'} B_{dn'} \right) \tag{7-29}$$

where the $J$'s are regarded as unknowns. Note the conceptual similarity between this approach and that used to solve the grey problem! In contrast to the integral operator approach described above, Feautrier's method is very easy to formulate and solve using the difference-equation procedures described in Chapter 6. In this method the "scattering" integral in equation (6–24) now involves the *entire* frequency spectrum. This shows explicitly the physically important fact that the radiation field at any frequency actually depends upon the field at *all other* frequencies. Using *current* estimate of $B_\nu$ and $\kappa^*_\nu$, equations (7-29) and the discretized form of (6-42) are solved for $J_\nu$ at all depths. These values are used in equation (7-28), which is solved for the new temperature that satisfies it (linearizing, in principle, both $B_\nu$ and $\kappa^*_\nu$ in terms of $\Delta T$ and iterating). Because opacities, etc., will be altered as a result of the changes in $T$, the whole process must be iterated to convergence.

In his original analysis Feautrier did not introduce the desired flux explicitly into the problem; one may do so easily, however, by using equation (7-27) to fix $|dT/dz|$ in the lower boundary condition [equation (6-44)]. If one uses Feautrier's method to solve the system, the cost is high because $N$, the number of frequencies, must be large (the angles can be eliminated in terms of variable Eddington factors); if radiative equilibrium is the *only* constraint involved, it is cheaper to use Rybicki's method, letting $\bar{J}$ denote the term in the numerator of (7-28); this equation in effect, replaces equation (6-48). Feautrier applied his scheme with good results for both LTE and non-LTE continuum models. The basic drawback of the method is that it is not clear how to generalize it, as it focuses entirely on the *temperature* correction (which is not sufficient in general).



*Exercise 7-6*: With the help of a computer, use the method just described to solve (a) the grey problem for $q(\tau)$; start with $q(\tau) = C$ and try several values for $C$; (b) the opacity-step problem of Exercise 7-5 [cf. (**128**; **605**; **38**)]. In part (a) use a *quadrature* for the frequency integral (not the exact results—this would make the problem trivial) and Rybicki's method for solving the final system.

An alternative method that also uses second-order difference equations for the transfer equation was proposed by Auer and Mihalas (**38**); this method is very easily generalized to extremely complicated problems, and forms the basis of the methods described in the final subsection of this section and in §7-5 for non-LTE problems. If the temperature structure $T^*(m)$ of the atmosphere were precisely that which produced radiative equilibrium, and $B^*_\nu$ the corresponding Planck function, then the solution of the transfer equation

$$\partial^2(f_\nu J_\nu)/\partial\tau_\nu{}^2 = J_\nu - B^*_\nu \tag{7-30}$$

with lower boundary condition [using equation (7-27)]

$$\frac{\partial(f_\nu J_\nu)}{\partial\tau_\nu} = \left(\frac{H}{\kappa^*_\nu}\frac{\partial B^*_\nu}{\partial T}\right) \Big/ \left(\int_0^\infty \frac{1}{\kappa^*_\nu}\frac{\partial B^*_\nu}{\partial T}\, d\nu\right) \tag{7-31}$$

would automatically satisfy the condition of radiative equilibrium

$$\sum_n w_n \kappa^*_{dn} J_{dn} = \sum_n w_n \kappa^*_{dn} B^*_{dn} \tag{7-32}$$

In practice we do not know $T^*(m)$ but only a current estimate $T(m)$; we therefore suppose that $T^*(m) = T(m) + \Delta T(m)$ and, ignoring changes in $\kappa^*_\nu$, expand $B^*_\nu = B_\nu(T) + (\partial B_\nu/\partial T)\,\Delta T$ and $(\partial B^*_\nu/\partial T) = (\partial B_\nu/\partial T) + (\partial^2 B_\nu/\partial T^2)\,\Delta T$ in equations (7-30) and (7-31), where $\Delta T$ must satisfy

$$\Delta T_d = \sum_n w_n \kappa^*_{dn}(J_{dn} - B_{dn}) \Big/ \sum w_n \kappa^*_{dn} (\partial B_n/\partial T)_d \tag{7-33}$$

We now can follow two possible approaches. We could expand equations (7-30) and (7-31) as described, eliminate $\Delta T$ using equation (7-33), and solve

the system

$$\left.\frac{\partial^2(f_n J_{dn})}{\partial \tau_n^{\,2}}\right|_d = J_{dn} - B_{dn} - \left(\frac{\partial B_n}{\partial T}\right)_d \frac{\sum_{n'} w_{n'}\kappa^*_{dn'}(J_{dn'} - B_{dn'})}{\sum_{n'} w_{n'}\kappa^*_{dn'}(\partial B_{n'}/\partial T)_d} \qquad (7\text{-}34)$$

with the corresponding boundary condition, regarding $J_{dn'}$ as unknown.

*Exercise 7-7*: Write out the perturbed boundary condition with $\Delta T$ eliminated.

Equation (7-34) closely resembles Feautrier's system (6-42) plus (7-29), and again involves the *entire* frequency spectrum; the same remarks about physical content apply. The solution of (7-34) satisfies both the transfer equation and the radiative equilibrium constraint (to first order) *simultaneously*. After the new intensities are found, the new temperature structure is evaluated using equation (7-33). Again if we wish to solve only for $\Delta T$, it is more efficient to use Rybicki's method, letting $\Delta T$ be the constraint variable replacing $\bar{J}$, and using equation (7-33) to replace (6-48) (this is the approach used for LTE atmospheres).

Both Feautrier's constraint procedure and the linearization method provide the equivalent of direct solution of the integral equation (7-21) but are simpler to implement than equation (7-24). Even though the linearization method introduces a local perturbation $\Delta T$, it defers knowledge of the mean intensity until the system is solved; it thus yields global convergence and is in no sense a lambda iteration. Further, the method is not inhibited by large opacity variations, for these enter only as coefficients of linear algebraic equations, and insofar as the correct inverses of these equations are obtained, we obtain the correct solution directly.

*Exercise 7-8*: Repeat Exercise 7-6 using the linearization method. Again check against the references in the literature, and use Rybicki's method to solve the system.

### A LINEARIZATION METHOD

Let us now draw together the various elements of the above discussion, and outline an efficient method for LTE model construction that experience shows to be general, stable, and effective (**35**; **275**). The basic thrust of the method is to write the system of transfer equations plus the constraints of hydrostatic and radiative equilibrium in terms of a *current* solution (which satisfies the constraints only imperfectly) and a *perturbation* of the fundamental variables ($T$, $N$) which, when evaluated, more nearly satisfies the constraints. In each equation we allow for the change produced in all variables by these perturbations, and for the coupling of these changes from one point in the atmosphere to another.

To begin, we need a starting solution for the structure of the atmosphere. We adopt a temperature distribution $T^4(\bar{\tau}_R) = \frac{3}{4}T^4_{\rm eff}[\bar{\tau}_R + q(\bar{\tau}_R)]$ on a



Rosseland-mean optical depth scale; we know that this is asymptotically the correct solution at depth. Here $q(\tau)$ may be the grey solution or some other function, based on previous results, which may differ from the grey value, particularly at the surface. We then integrate the hydrostatic equation (7-8) in the approximate form

$$(dp_g/dm) = g - (\sigma_R T^4_{\rm eff}/c)(\bar{\chi}_R/\rho) \qquad (7\text{-}35)$$

simultaneously with the definition of the optical-depth scale

$$d\bar{\tau}_R = (\bar{\chi}_R/\rho)\, dm \qquad (7\text{-}36)$$

step-by-step on a mesh $\{m_d\}$, $(d = 1, \ldots, D)$. This yields $(N_d, T_d)$ at each point on the mesh, and using the method described in §5-2 we solve the LTE equation of state for $n_e$ and $n_i^*(N, T)$ for all atomic and ionic levels. We then calculate $\chi^*_{dn}$ and $\eta^*_{dn}$ from equations (7-3) and (7-4) on a frequency mesh $\{\nu_n\}$, and evaluate the mean intensities $J_{dn}$ and Eddington factors $f_{dn}$ from a formal solution of the transfer equation (6-30) with a *given* source function $S_{dn}$.

The starting solution provides enough information to evaluate the radiation pressure gradient in equations (7-9) and (7-10), which we can therefore integrate to find a new estimate of the run of the total number density $N_d$, and new LTE occupation numbers $n_i^*$. Further, knowing the variable Eddington factors we can solve the discrete form of the transfer equations (6-42) through (6-44), which in light of (7-27) we can write as

$$(f_{2n}J_{2n} - f_{1n}J_{1n})/\Delta\tau_{\frac{3}{2},n} = h_n J_{1n} \qquad (7\text{-}37a)$$

$$\frac{f_{d-1,n}J_{d-1,n}}{\Delta\tau_{d-\frac{1}{2},n}\,\Delta\tau_{dn}} - \frac{f_{dn}}{\Delta\tau_{dn}}\left(\frac{1}{\Delta\tau_{d-\frac{1}{2},n}} + \frac{1}{\Delta\tau_{d+\frac{1}{2},n}}\right) J_{dn} + \frac{f_{d+1,n}J_{d+1,n}}{\Delta\tau_{d+\frac{1}{2},n}\,\Delta\tau_{dn}}$$
$$= \left(1 - \frac{n_{e,d}\sigma_e}{\chi_{dn}}\right) J_{dn} - \frac{\eta_{dn}}{\chi_{dn}}, \quad (d = 2, \ldots, D-1) \qquad (7\text{-}37b)$$

and (at $d = D$),

$$(f_{Dn}J_{Dn} - f_{D-1,n}J_{D-1,n})/\Delta\tau_{D-\frac{1}{2},n}$$
$$= H\chi_{Dn}^{-1}(\partial B_\nu/\partial T)_{Dn}/\sum_n w_n[\chi_{Dn}^{-1}(\partial B_\nu/\partial T)_{Dn}] \qquad (7\text{-}37c)$$

where
$$\Delta\tau_{d\pm\frac{1}{2},n} \equiv \frac{1}{2}\left[(\chi_{d\pm1,n}/\rho_{d\pm1}) + (\chi_{dn}/\rho_{dn})\right]|m_{d\pm1} - m_d| \qquad (7\text{-}38a)$$

$$\Delta\tau_{dn} \equiv \frac{1}{2}(\Delta\tau_{d-\frac{1}{2},n} + \Delta\tau_{d+\frac{1}{2},n}) \qquad (7\text{-}38b)$$

First-order boundary conditions have been written for simplicity; it is easy to include second-order terms.

If we solved equations (7-37) we would find that the constraint of radiative equilibrium, equation (7-13), is not satisfied; we must therefore change the temperature $T(m)$ in such a way as to more accurately satisfy the conditions of radiative equilibrium, and iterate. There are two difficulties: (a) the problem is *nonlinear*, and (b) the coupling is *global*. That is, any change $\delta T_d$ implies a change $\delta N_d$ (from hydrostatic equilibrium) and therefore $\delta\chi_d$, $\delta\eta_d$, and hence $\delta J_{d'n}$ at *all* $d'$ and $n$ *throughout the atmosphere*. To handle these problems, we *linearize* the equations, replacing each variable $x$ by $x_0 + \delta x$, and retain only first-order terms in the $\delta$'s. The power of this method is that (a) it may be applied to a wide variety of constraints, and (b) it produces systems that allow for the effects of a change in a variable at a given point in the atmosphere upon all other variables at all other points. In particular, the linearized transfer equations describe fully how a change in material properties or radiation field at any point propagates and affects the solution at every other point. We may use the linearized transfer equations frequency-by-frequency to eliminate the $\delta J$'s from the constraint equations (radiative and hydrostatic equilibrium), yielding a final system for the perturbations of the "fundamental" variables $\delta N$ and $\delta T$. Thus, linearizing the transfer equation (assuming the Eddington factors remain unchanged) we have, away from the boundaries, and at each frequency $\nu_n$,

$$\frac{f_{d-1,n}\,\delta J_{d-1,n}}{\Delta\tau_{d-\frac{1}{2},n}\,\Delta\tau_{dn}} - \left[\frac{f_{dn}}{\Delta\tau_{dn}}\left(\frac{1}{\Delta\tau_{d-\frac{1}{2},n}} + \frac{1}{\Delta\tau_{d+\frac{1}{2},n}}\right) + \left(1 - \frac{n_{e,d}\sigma_e}{\chi_{dn}}\right)\right]\delta J_{dn}$$
$$+ \frac{f_{d+1,n}\,\delta J_{d+1,n}}{\Delta\tau_{d+\frac{1}{2},n}\,\Delta\tau_{dn}} + a_{dn}\,\delta\omega_{d-1,n} + b_{dn}\,\delta\omega_{dn} + c_{dn}\,\delta\omega_{d+1,n}$$
$$- (\eta_{dn} + n_{e,d}\sigma_e J_d)\frac{\delta\chi_{dn}}{\chi_{dn}^2} + \frac{\delta\eta_{dn}}{\chi_{dn}} + \frac{\sigma_e J_{dn}}{\chi_{dn}}\,\delta n_{e,d}$$
$$= \beta_{dn} + J_{dn} - (n_{e,d}\sigma_e J_{dn} + \eta_{dn})/\chi_{dn} \tag{7-39}$$

where
$$\alpha_{dn} \equiv (f_{dn}J_{dn} - f_{d-1,n}J_{d-1,n})/(\Delta\tau_{d-\frac{1}{2},n}\,\Delta\tau_{dn}) \tag{7-40}$$
$$\gamma_{dn} \equiv (f_{dn}J_{dn} - f_{d+1,n}J_{d+1,n})/(\Delta\tau_{d+\frac{1}{2},n}\,\Delta\tau_{dn}) \tag{7-41}$$
$$\beta_{dn} \equiv \alpha_{dn} + \gamma_{dn} \tag{7-42}$$
$$a_{dn} \equiv \left[\alpha_{dn} + \frac{1}{2}\beta_{dn}(\Delta\tau_{d-\frac{1}{2},n}/\Delta\tau_{dn})\right]\Big/(\omega_{d-1,n} + \omega_{dn}) \tag{7-43}$$
$$c_{dn} \equiv \left[\gamma_{dn} + \frac{1}{2}\beta_{dn}(\Delta\tau_{d+\frac{1}{2},n}/\Delta\tau_{dn})\right]\Big/(\omega_{dn} + \omega_{d+1,n}) \tag{7-44}$$
$$b_{dn} \equiv a_{dn} + c_{dn} \tag{7-45}$$

and
$$\omega_{dn} \equiv \chi_{dn}/\rho_d \tag{7-46}$$



Note that equations (7-39) through (7-46) apply both for LTE and non-LTE cases.

*Exercise 7-9*: (a) Derive equation (7-39). (b) Derive linearized expressions for the upper and lower boundary conditions, equations (7-37a) and (7-37c). See also (**437**).

In equation (7-39), *assuming LTE*, all material variations are expressed in terms of $\delta N$ and $\delta T$. Thus, from equations (7-6) and (7-3)

$$\delta\rho_d = \bar{m}(\delta N_d - \delta n_{e,d}) \tag{7-47}$$
$$\delta\chi_{dn}^* = (\partial\chi_n^*/\partial T)_d\,\delta T_d + (\partial\chi_n^*/\partial n_e)_d\,\delta n_{e,d} + \sum_i (\partial\chi_n^*/\partial n_i^*)_d\,\delta n_{i,d}^* \tag{7-48}$$

with a similar expression for $\delta\eta_{nd}^*$; in equation (7-48), $\partial/\partial T$ applies to the *explicit* appearances of $T$ in $\exp(-h\nu/kT)$, $\alpha_{\kappa\kappa}(\nu, T)$, etc., and similarly for $\partial/\partial n_e$. Now from equation (5-35) we have relations of the form

$$\delta n_{i,d}^* = (\partial n_i^*/\partial T)_N\big|_d\,\delta T_d + (\partial n_i^*/\partial N)_T\big|_d\,\delta N_d \tag{7-49}$$

and similarly for $\delta n_{e,d}$, so that the linearized values of $\rho_d$, $\chi_{dn}^*$ and $\eta_{dn}^*$ can all be collapsed down to expressions of the form

$$\delta\chi_{dn}^* = \left(\frac{\partial\chi_n^*}{\partial T}\right)_N\Big|_d\,\delta T_d + \left(\frac{\partial\chi_n^*}{\partial N}\right)_T\Big|_d\,\delta N_d \tag{7-50}$$

The end result is that equation (7-39) reduces to a formula involving the perturbations at three adjacent points $(d-1, d, d+1)$, of the general form

$$\sum_{d'=d-1}^{d+1} T_{dd',n}\,\delta J_{d'n} + \sum_{d'=d-1}^{d+1} U_{dd',n}\,\delta N_{d'} + \sum_{d'=d-1}^{d+1} V_{dd',n}\,\delta T_{d'} = K_d \tag{7-51}$$

Similarly, the linearized constraint of radiative equilibrium [equation (7-13)] is

$$\sum_n w_n(\chi_{dn} - n_{e,d}\sigma_e)\,\delta J_{dn} + \sum_n w_n(J_{dn}\,\delta\chi_{dn} - \delta\eta_{dn} - \sigma_e J_{dn}\,\delta n_{e,d})$$
$$= \sum_n w_n[\eta_{dn} - (\chi_{dn} - n_{e,d}\sigma_e)J_{dn}] \tag{7-52}$$

and hydrostatic equilibrium [equation (7-9)] yields

$$(4\pi/c)\sum_n w_n(f_{dn}\,\delta J_{dn} - f_{d-1,n}\,\delta J_{d-1,n})$$
$$+ k(T_d\,\delta N_d + N_d\,\delta T_d - T_{d-1}\,\delta N_{d-1} - N_{d-1}\,\delta T_{d-1})$$
$$= g(m_d - m_{d-1}) - N_d kT_d + N_{d-1}kT_{d-1}$$
$$- (4\pi/c)\sum_n w_n(f_{dn}J_{dn} - f_{d-1,n}J_{d-1,n}) \tag{7-53}$$

In (7-52) $\delta\chi$ and $\delta\eta$ are replaced with expressions of the form (7-50). Equation (7-52) involves information only at one depth-point while (7-53) involves two.

*Exercise 7-10*: (a) Verify equations (7-52) and (7-53). (b) Linearize the upper boundary condition on hydrostatic equilibrium.

The whole system, for all depths and frequencies, can be organized into a form suitable for a Rybicki-method solution. Thus let

$$\boldsymbol{\delta J}_n \equiv (\delta J_{1n}, \delta J_{2n}, \ldots, \delta J_{Dn})^T, \quad (n = 1, \ldots, N) \tag{7-54}$$

$$\boldsymbol{\delta T} \equiv (\delta T_1, \delta T_2, \ldots, \delta T_D)^T \tag{7-55}$$

$$\boldsymbol{\delta N} \equiv (\delta N_1, \delta N_2, \ldots, \delta N_D)^T \tag{7-56}$$

Then equations (7-39), (7-52) and (7-53) yield

$$\begin{pmatrix} \mathbf{T}_1 & 0 & \cdots & 0 & \mathbf{U}_1 & \mathbf{V}_1 \\ 0 & \mathbf{T}_2 & & \cdot & \mathbf{U}_2 & \mathbf{V}_2 \\ \cdot & & \ddots & \cdot & \cdot & \cdot \\ \cdot & & & 0 & \cdot & \cdot \\ 0 & & \cdots & \mathbf{T}_N & \mathbf{U}_N & \mathbf{V}_N \\ \mathbf{W}_1 & \mathbf{W}_2 & \cdots & \mathbf{W}_N & \mathbf{A} & \mathbf{B} \\ \mathbf{X}_1 & \mathbf{X}_2 & \cdots & \mathbf{X}_N & \mathbf{C} & \mathbf{D} \end{pmatrix} \begin{pmatrix} \boldsymbol{\delta J}_1 \\ \boldsymbol{\delta J}_2 \\ \cdot \\ \cdot \\ \boldsymbol{\delta J}_N \\ \boldsymbol{\delta N} \\ \boldsymbol{\delta T} \end{pmatrix} = \begin{pmatrix} \mathbf{K}_1 \\ \mathbf{K}_2 \\ \cdot \\ \cdot \\ \mathbf{K}_N \\ \mathbf{L} \\ \mathbf{M} \end{pmatrix} \tag{7-57}$$

Each "element" in equation (7-57) is a matrix of dimension $(D \times D)$; the first $N$ "rows" represent transfer equations, the next-to-last "row" represents radiative equilibrium, and the last, hydrostatic equilibrium. The **T**, **U**, and **V** matrices are tridiagonal; **W**, **A**, and **B** are diagonal; **X**, **C**, and **D** are bidiagonal. The vectors **K**, **L**, and **M** give the *errors* in the transfer and constraint equations arising with the current estimates of the radiation field, temperature, and density. Equation (7-57) is solved by performing eliminations of $\boldsymbol{\delta J}_n$ from the $n$th "row" into the last two "rows," one frequency at a time. Thus we solve for

$$\boldsymbol{\delta J}_n + (\mathbf{T}_n^{-1}\mathbf{U}_n)\,\boldsymbol{\delta N} + (\mathbf{T}_n^{-1}\mathbf{V}_n)\,\boldsymbol{\delta T} = \mathbf{T}_n^{-1}\mathbf{K}_n \tag{7-58}$$

and eliminate $\boldsymbol{\delta J}_n$ to obtain a final system of the form

$$\begin{pmatrix} \mathbf{P} & \mathbf{Q} \\ \mathbf{R} & \mathbf{S} \end{pmatrix} \begin{pmatrix} \boldsymbol{\delta N} \\ \boldsymbol{\delta T} \end{pmatrix} = \begin{pmatrix} \mathbf{F} \\ \mathbf{G} \end{pmatrix} \tag{7-59}$$



where $\mathbf{P} \equiv \mathbf{A} - \sum_n \mathbf{W}_n\mathbf{T}_n^{-1}\mathbf{U}_n$; $\mathbf{Q} \equiv \mathbf{B} - \sum_n \mathbf{W}_n\mathbf{T}_n^{-1}\mathbf{V}_n$

$\mathbf{R} \equiv \mathbf{C} - \sum_n \mathbf{X}_n\mathbf{T}_n^{-1}\mathbf{U}_n$; $\mathbf{S} \equiv \mathbf{D} - \sum_n \mathbf{X}_n\mathbf{T}_n^{-1}\mathbf{V}_n$

$\mathbf{F} \equiv \mathbf{L} - \sum_n \mathbf{W}_n\mathbf{T}_n^{-1}\mathbf{K}_n$; and $\mathbf{G} \equiv \mathbf{M} - \sum_n \mathbf{X}_n\mathbf{T}_n^{-1}\mathbf{K}_n$

The final system (7-59) is solved for $\boldsymbol{\delta N}$ and $\boldsymbol{\delta T}$.

Using $\boldsymbol{\delta N}$ and $\boldsymbol{\delta T}$ to revise the density and temperature, we can at each point $m_d$ solve for new values of all the $n_{i,d}^*$'s and $n_{e,d}$ [using equations (5-27) through (5-31) iterated to consistency] and hence new values of $\chi_{dn}^*$ and $\eta_{dn}^*$, which are, in turn, used to obtain a formal solution of the transfer equation for new values of $J_{dn}$ and $f_{dn}$ $(d = 1, \ldots, D;\ n = 1, \ldots, N)$. We use these new estimates to reconstruct equation (7-57), and iterate; as the solution improves, $\mathbf{K}_n$, **L**, and **M** all $\to 0$, hence $\boldsymbol{\delta N}$ and $\boldsymbol{\delta T} \to 0$.

The computing time per iteration scales with $N$ and $D$ (the number of frequencies and depths) as $T = c(2N)D^2 + c'(2D)^3$, which is *linear* in $N$ (so that *many* frequencies may be included—e.g., for line-blanketing). Actually, experience has shown (**35**; **275**) that the solution can be greatly economized in most cases by assuming that the gas pressure $p_g$ will remain unchanged during the linearization [as it will if the radiation-pressure terms in (7-9) are negligibly small]. We then rewrite all expansions as $\delta x = (\partial x/\partial T)_{p_g}\,\delta T + (\partial x/\partial p_g)_T\,\delta p_g$, where

$$(\partial x/\partial T)_{p_g} \equiv (\partial x/\partial T)_N + (\partial x/\partial N)_T(\partial N/\partial T)_{p_g},$$

and then explicitly assume $\delta p_g \equiv 0$. This eliminates the last "row" of system (7-57), and we solve only for $\boldsymbol{\delta T}$; the computing time then becomes $T = cND^2 + c'D^3$. The method just described has not yet been widely used, but its advantages are manifest; it is likely to be the preferred method in future work on LTE model atmospheres.

## 7-3 Convection and Models for Late-Type Stars

The energy transport in a stellar atmosphere may proceed by radiative transfer or by convection; the process that actually occurs is that which is more efficient. In general terms, radiative equilibrium prevails in spectral types A and earlier, while convection becomes important in the middle F-stars and dominates in later types. The convective flow in stellar atmospheres is *turbulent* [see, e.g., (**90**)] and consists of a complicated hierarchy of "eddies" or "bubbles" moving and interacting in an extremely involved way. The situation poses many physical and mathematical problems of

great complexity, and a definitive convection theory does not yet exist. We shall, therefore, consider only the phenomenological *mixing-length theory*, which contains a number of the basic physical ingredients and provides a framework for at least illustrating the effects of convection.

### THE SCHWARZSCHILD STABILITY CRITERION

Suppose we have an atmosphere in radiative equilibrium. We then ask whether an element of material, when displaced from its original position, experiences forces that tend to move it farther in the direction of its displacement. If so, the atmosphere is unstable against mass motions, and convection will occur; if not, the motion will be damped and will die out, and radiative equilibrium will persist. The basic criterion for stability against convection was established by K. Schwarzschild (**416**, 25).

Consider a small element of gas whose position is perturbed upward by a distance $\Delta r$ in the atmosphere. We suppose that (a) the movement is so slow that the element remains in pressure equilibrium with its surroundings and (b) the element does not exchange energy with its surroundings (i.e., the process is *adiabatic*). Because the pressure drops as the element rises, the gas expands, and the density decreases by an amount $(\Delta\rho)_E = (d\rho/dr)_A\,\Delta r$; the subscript E denotes "element" and A denotes "adiabatic." If, at its new position, the density of the element is less than that of its surroundings, it experiences a buoyancy force and will continue to rise. That is, if $(d\rho/dr)_R$ is the density gradient in the radiative surroundings, instability occurs if

$$(\Delta\rho)_E = (d\rho/dr)_A\,\Delta r < (\Delta\rho)_R = (d\rho/dr)_R\,\Delta r \tag{7-60}$$

(recall that $d\rho/dr < 0$). We may write equation (7-60) in a more convenient form. In the adiabatic element (which we shall momentarily assume is a perfect gas) the equation of state is $\ln p = \gamma \ln \rho + C$, while in the radiative surroundings (again assumed to be a perfect gas) $\ln p = \ln \rho + \ln T + C'$. Using these relations to compute $(d\rho/dr)_A$ and $(d\rho/dr)_R$, and demanding the pressure gradients be equal, we find from equation (7-60) that the Schwarzschild condition for instability is

$$[(\gamma - 1)/\gamma](-d \ln p/dr)_R < (-d \ln T/dr)_R \tag{7-61}$$

or
$$\nabla_R \equiv (d \ln T/d \ln p)_R > (\gamma - 1)/\gamma = (d \ln T/d \ln p)_A \equiv \nabla_A \tag{7-62}$$

In stellar atmospheres the gas is not perfect because of the effects of ionization and radiation pressure; we may account for this by generalizing $\gamma$ to $\Gamma$ (**160**, 57) and writing $\nabla_A = (\Gamma - 1)/\Gamma$ where $\Gamma$ will not, in general, equal its value for a perfect monatomic gas, namely $\gamma = (C_p/C_v) = \frac{5}{3}$. Convenient formulae for the calculation of $\Gamma$, allowing for radiation pressure and ionization, have been given by several authors [see, e.g., (**638**, §56; **643**; **364**)]. These effects can be of major importance, and may drastically lower $\nabla_A$, and hence the critical value of $\nabla_R$ at which convection occurs. Thus for a perfect monatomic gas $\nabla_A = \frac{2}{5}/\frac{5}{3} = 0.4$, while for pure radiation pressure $\Gamma = \frac{4}{3}$, so $\nabla_A = 0.25$, and for conditions where hydrogen is ionizing $\Gamma$ may be only 1.1 so $\nabla_A$ drops to 0.1! These results clearly suggest that *we may expect convection to occur in hydrogen ionization zones.* This expectation is strengthened by noting that in the limit of the diffusion approximation $(-dT/dr) = (3\pi F\bar{\chi}_R)/(16\sigma_R T^3)$, which implies (from hydrostatic equilibrium) $\nabla_R = (3\pi F\bar{\chi}_R p)/(16\sigma_R g\rho T^4)$. From this we see that large values of the opacity require that the radiative gradient must be steep in order to drive the flux $F$ through the atmosphere. The opacity of stellar material becomes large when hydrogen is appreciably excited into its upper states; this happens at about the same conditions where ionization occurs and causes $\Gamma$ to decrease. The two effects work together and imply that the radiative gradient does, in fact, exceed the adiabatic gradient in the hydrogen ionization zone, so that convection occurs. The importance of these mechanisms and the existence of extensive hydrogen convection zones in stellar envelopes was first recognized by Unsöld (**636**).



In the earliest-type stars, hydrogen is essentially completely ionized throughout the envelope, and radiative equilibrium prevails (thin, weak convection zones associated with $He^0$ and $He^+$ ionization exist, but transport only a tiny fraction of the flux). In the A-stars, thin hydrogen convection zones begin to develop at shallow depths ($\tau \approx 0.2$). In the F-stars, the convection zone starts somewhat deeper, and becomes thicker; by types F2 to F5 convection will transport essentially all of the flux at some point within the zone. For later and later types the zone extends ever deeper, and convection becomes more efficient; in the M-stars the convective envelope is so extensive that it determines the structure of the star as a whole (**396**).

### MIXING-LENGTH THEORY

The basic physical picture used in the mixing-length theory is that the transport in the unstable layer is effected by turbulent elements moving upward and downward through a surrounding environment. The upward (downward) moving elements have an excess (deficiency) of thermal energy relative to the surrounding material. At the end of some characteristic distance, the *mixing-length*, one supposes that these elements "dissolve" smoothly into the surroundings, delivering any excess energy they possess or absorbing any deficiency. A direct energy transport results, and the temperature gradient is decreased below that which would occur if the only transport mechanism were radiation. To characterize the process we introduce the following gradients: $\nabla_R$ is the radiative gradient that would occur *if convection were suppressed*; $\nabla_A$ is the adiabatic gradient; $\nabla_E$ is the gradient of the convective elements; and $\nabla$ is the "true" gradient of the surroundings in the final state where both radiation and convection transport the total

flux together. In general we will have

$$\nabla_R \geqslant \nabla \geqslant \nabla_E \geqslant \nabla_A \tag{7-63}$$

Consider now a rising element of material. If $\delta T$ is the temperature difference between the element and its surroundings, the excess energy delivered per unit volume when the element merges into the surroundings is $\rho C_p\, \delta T$. The temperature difference arises from the difference between the gradients of the element and the surroundings. Thus for elements traveling over a distance $\Delta r$ with an average speed $\bar{v}$, the energy flux transported is

$$\pi F_{\text{conv}} = \rho C_p \bar{v} \delta T = \rho C_p \bar{v}[(-dT/dr) - (-dT/dr)_E]\, \Delta r \tag{7-64}$$

At a given level in the atmosphere we will find elements distributed at random over their paths of travel; averaging over all elements, we set $\Delta r = l/2$ where $l$ is the mixing length. Further, using the hydrostatic equation $(dp/dr) = -\rho g$, and introducing the *pressure scale height* $H \equiv (-d \ln p/dr)^{-1} = p/(g\rho)$ we can rewrite (7-64) as

$$\pi F_{\text{conv}} = \tfrac{1}{2} \rho C_p \bar{v} T(\nabla - \nabla_E)(l/H) \tag{7-65}$$

To estimate $\bar{v}$, we calculate the work done by buoyant forces on an element and equate this to its kinetic energy. If $\delta\rho$ is the density difference between the element and its surroundings, then the buoyant force is $f_b = -g\, \delta\rho$. The equation of state yields $\ln \rho = \ln p - \ln T + \ln \mu$, where $\mu$ is now considered to be *variable* to allow for effects of ionization and radiation pressure. Thus we may write $d(\ln \rho) = d(\ln p) - Q d(\ln T)$, where $Q \equiv 1 - (\partial \ln \mu/\partial \ln T)_p$, and, demanding pressure equilibrium ($\delta p = 0$), we have $\delta\rho = -Q\rho\, \delta T/T$, so that

$$f_b = (gQ\rho/T)\, \delta T = (gQ\rho/T)[(-dT/dr) - (-dT/dr)_E]\, \Delta r \tag{7-66}$$

The buoyancy force is thus linear in the displacement; integrating over a total displacement $\Delta$, and setting $\Delta = l/2$ to account for the average over all elements passing the point under consideration, we obtain the average work done on the elements

$$\bar{w} = \int_0^{\Delta} f_b(\Delta r)\, d(\Delta r) = (gQ\rho H/8)(\nabla - \nabla_E)(l/H)^2 \tag{7-67}$$

We now suppose that about one-half of this work will be lost to "friction" in pushing aside other turbulent elements and the other half will provide the kinetic energy of the element (i.e., $\frac{1}{2}\rho\bar{v}^2 \approx \frac{1}{2}\bar{w}$) from which we find

$$\bar{v} = (gQH/8)^{\frac{1}{2}}(\nabla - \nabla_E)^{\frac{1}{2}}(l/H) \tag{7-68}$$

and, therefore, from equation (7-65),

$$\pi F_{\text{conv}} = (gQH/32)^{\frac{1}{2}}(\rho C_p T)(\nabla - \nabla_E)^{\frac{3}{2}}(l/H)^2 \tag{7-69}$$



One of the uncertaintities of this approach lies in the question of how to specify the mixing-length $l$; usually it is assumed that $l$ is simply some multiple of $H$, say 1 or 2.

To complete the theory, we need another relation that will allow us to express $\nabla$ and $\nabla_E$ in terms of $\nabla_R$ and $\nabla_A$; this may be done, following Unsöld, by considering the efficiency of the convective transport. As an element rises, its temperature exceeds that of the surroundings (which accounts for the energy transport); the temperature excess implies that it will lose some energy to its surroundings by radiation. This energy loss will diminish the excess energy content of the element and therefore decrease the energy yield when the element "dissolves" at the end of its mixing length. We therefore define the efficiency parameter as

$$\gamma = \frac{\text{excess energy content at time of dissolution}}{\text{energy lost by radiation during lifetime of element}} \tag{7-70}$$

The excess energy content of the element is proportional to $(\nabla - \nabla_E)$ [cf. equation (7-65)]; had the element moved adiabatically, the energy content would have been proportional to $(\nabla - \nabla_A)$. Therefore the loss by radiation is proportional to $(\nabla - \nabla_A) - (\nabla - \nabla_E) = (\nabla_E - \nabla_A)$ so that

$$\gamma = (\nabla - \nabla_E)/(\nabla_E - \nabla_A) \tag{7-71}$$

Alternatively, we may calculate the quantities in the numerator and denominator of equation (7-70) in terms of local variables. Thus for an element of volume $V$, with excess temperature $\delta T$, the excess energy content is $\rho C_p V\, \delta T$. The radiative loss depends on whether the element is optically thin or thick. In the thin limit the rate of energy loss will be $4\pi\bar{\chi}_R\, \Delta B$, from a volume $V$, over a lifetime $(l/\bar{v})$. Assuming an average excess of $(\delta T/2)$ over this lifetime, we have

$$\begin{aligned} \gamma_{\text{thin}} &= (\rho C_p V\, \delta T)/[4\pi(4\sigma_R T^3/\pi)(\delta T/2)(\bar{\chi}_R V)(l/\bar{v})] \\ &= (\rho C_p \bar{v})/(8\sigma_R T^3 \tau_e) \end{aligned} \tag{7-72}$$

where $\tau_e$ denotes the optical thickness of the characteristic element size $l$, $\tau_e = \bar{\chi}_R l$. Equation (7-72) applies when $\tau_e \ll 1$. At the opposite extreme, $\tau_e \gg 1$, we apply the diffusion approximation to determine the radiative flux lost by an element of characteristic size $l$, with fluctuation $\delta T$, by writing $(-dT/dr) \approx (\delta T/l)$. Assuming a surface area A and the same lifetime as before, we now have

$$\gamma_{\text{thick}} = (\rho C_p V\, \delta T)/[(16\sigma_R T^3/3\bar{\chi}_R)(\delta T/l)A(l/\bar{v})] = (\rho C_p \bar{v}/16\sigma_R T^3)3\bar{\chi}_R(V/A) \tag{7-73}$$

The choice of $(V/A)$ is ambiguous and introduces another source of uncertainty into the theory; if the elements are presumed to be spherical, $(V/A) = l/3$ and

$$\gamma_{\text{thick}} = \tfrac{1}{2}\tau_e(\rho C_p \bar{v})/(8\sigma_R T^3) \tag{7-74}$$

We interpolate between the two extreme cases by writing

$$\gamma = [(\rho C_p \bar{v})/(8\sigma_R T^3)] \cdot [(1 + \tfrac{1}{2}\tau_e{}^2)/\tau_e] \tag{7-75}$$

Combining equations (7-71) and (7-75), and substituting equation (7-68) for $\bar{v}$ we derive finally

$$\frac{\nabla_E - \nabla_A}{(\nabla - \nabla_E)^{\frac{1}{2}}} = \frac{16\sqrt{2}\,\sigma_R T^3}{\rho C_p (gQH)^{\frac{1}{2}}(l/H)} \frac{\tau_e}{(1 + \frac{1}{2}\tau_e{}^2)} \equiv B \tag{7-76}$$

The final requirement we place upon the theory is that the correct total flux be transported by radiation and convection together—i.e.,

$$\pi F = \pi F_{\text{rad}} + \pi F_{\text{conv}} = \sigma_R T_{\text{eff}}^4 \tag{7-77}$$

The mixing-length theory described above is the simplest (and most widely used!) convection theory in astrophysics. Numerous refinements have been proposed, attempting to introduce nonlocal information into the theory; it would take us too far afield to attempt to describe these here, and the interested reader should examine the literature. [See, e.g., (**594**; **595**; **450**, 237; **479**) and the references cited therein.]

### CONVECTIVE MODEL ATMOSPHERES

The computation of convective model atmospheres is more complicated than for radiative models (even assuming the mixing-length theory) because there are two transport mechanisms that must be brought into a final balance to satisfy equation (7-77). We may proceed as follows. Suppose we assume some specification of the temperature distribution—e.g., the grey distribution on a Rosseland-mean optical-depth scale. We then carry out a step-by-step integration of equations (7-35) and (7-36), as before. At each point we may calculate $\nabla_R = \nabla_R(T, p, p_g)$ and $\nabla_A = \nabla_A(T, p, p_g)$. If at some point we find that the instability criterion is satisfied, we must determine the true gradient $\nabla$, $\nabla_R \geqslant \nabla \geqslant \nabla_A$, which satisfies equation (7-77). If the instability occurs deep enough for the diffusion approximation to be valid, then $(F_{\text{rad}}/F) = (\nabla/\nabla_R)$, and equations (7-77) and (7-69) reduce to

$$A(\nabla - \nabla_E)^{\frac{3}{2}} = \nabla_R - \nabla \tag{7-78}$$



where $A$ depends only on local variables. Adding $(\nabla - \nabla_E) + (\nabla_E - \nabla_A)$ to both sides of equation (7-78) and using equation (7-76) to eliminate $(\nabla_E - \nabla_A)$, we find a cubic equation for $(\nabla - \nabla_E)^{\frac{1}{2}} \equiv x$, namely

$$A(\nabla - \nabla_E)^{\frac{3}{2}} + (\nabla - \nabla_E) + B(\nabla - \nabla_E)^{\frac{1}{2}} = (\nabla_R - \nabla_A) \tag{7-79}$$

which may be solved by standard methods for the root $x_0$. We thus obtain the true gradient $\nabla = \nabla_A + Bx_0 + x_0{}^2$, and proceed with the integration, now regarding $T$ as a function of $p$. If, at some point, convection ceases, we revert to the original $T(\bar{\tau}_R)$ relation (adjusted to match the current values of $T$ and $\bar{\tau}_R$) and continue the integration into a radiative zone.

In the case that the material is presumed grey [or, for nongrey material, the convection zone is really deep enough that the diffusion approximation is correct, *and* the true nongrey temperature distribution is known near the surface] the treatment described above is essentially exact. Using this approach for grey atmospheres, Vitense (**653**) performed computations for a wide range of effective temperatures and gravities; this work nicely delineates the role of convection in stellar atmospheres over much of the H–R diagram. In a general way, the outermost layers can always be expected to be in radiative equilibrium because densities and opacities are small and radiative transport is more efficient than convective. In deeper layers, the opacity and density rise, ionization may occur, and convection may begin. Convection will have its largest effects in stars of low effective temperatures (in which the hydrogen is essentially neutral in the outer layers) and high gravities (which imply large densities and heat capacity, hence efficient thermal transport). When convection is efficient, it will transport practically all the flux, and $\nabla$ will be close to $\nabla_A$; indeed, in stellar interiors, convection (which it occurs) is so efficient that one may set $\nabla \equiv \nabla_A$ and dispense with the mixing-length theory entirely. When convection is inefficient, the true gradient $\nabla$ will lie close to $\nabla_R$, and a substantial part of the flux may be carried by radiation; in this regime the uncertainties of the mixing-length theory make themselves felt fully.

When the convection zone lies close enough to the surface that the diffusion approximation used to derive equation (7-78) is invalid, it is then necessary to calculate $F_{\text{rad}}$ from the solution of the nongrey transfer equation, and employ an iterative temperature-correction procedure. In any such procedure it is essential to account for changes in both $F_{\text{rad}}$ and $F_{\text{conv}}$ induced by the proposed alteration of the temperature structure. Methods for constructing convective models based on a generalization of the Avrett–Krook procedure have been used to study F-type main-sequence stars (**422**), middle-type supergiants (**500**), and M-stars (dwarfs through supergiants) (**48**). A detailed description of a computer code that treats convection is given in (**379**). An extensive grid of nongrey models ($4000°\text{K} \leqslant T_{\text{eff}} \leqslant 50{,}000°\text{K}$, $2 \leqslant \log g \leqslant$

5), including convection effects where appropriate and making allowance for line-blanketing, is available (**247**, 377). More limited grids of blanketed convective models may be found in (**512**; **513**; **514**), and extensive computations for M giants and supergiants, allowing for molecular line-blanketing, are given in (**341**; **342**). The solar convection zone has been studied with both the mixing-length approximation (**652**) [see also (**479**)] and more detailed hydrodynamical theories (**99**; **100**). Recently, methods for computing convective models using a linearization procedure similar to that described in §7-2 have been developed (**274**; **275**; **479**). The basic change in the formulation is to use equation (7-77) as the energy balance equation; introducing a discrete representation of the flux [cf. equations (6-15) and (6-26)]. On an angle-frequency mesh $\{\mu_i, \nu_i\}$ we can write

$$4\pi \sum_{i=1}^{J} w_i \mu_i^2 (u_{d+1,i} - u_{di})/\Delta\tau_{d+\frac{1}{2},i} + \pi F_{\text{conv}}|_{d+\frac{1}{2}} = \sigma_R T_{\text{eff}}^4 \qquad (7\text{-}80)$$

The convective flux can be regarded as $F_{\text{conv}}(p, p_g, T, \nabla)$ [given these variables, $\nabla_E$ follows from equation (7-76) and $F_{\text{conv}}$ from equation (7-69)]. The radiative term may be linearized as before. In linearizing the convective term, the total pressure is fixed, and the derivatives appearing in the expression

$$F_{\text{conv}} = F_{\text{conv}}^0 + (\partial F_c/\partial p_g)\,\delta p_g + (\partial F_c/\partial T)\,\delta T + (\partial F_c/\partial \nabla)\,\delta\nabla \qquad (7\text{-}81)$$

are computed numerically. Several approximations are introduced (**274**) to reduce this to an expression in $\delta T$ only, and practical procedures for handling numerical problems have been developed (**274**; **275**). Improvements in convergence might be obtained by including terms in $\delta N$ as well as $\delta T$, but, as described earlier, this is inherently more costly.

At the present time, convection theory as applied in stellar atmospheres analysis is only heuristic; improvements to the physical theory are being actively pursued and, when more accurate treatments of convection become available, our understanding of the atmospheres of late-type stars will be improved significantly.

## 7-4 Results of LTE Model-Atmosphere Calculations for Early-Type Stars

The largest group of reliable model atmospheres available pertains to solar and earlier spectral types; therefore we shall confine attention primarily to these stars. For later types, many difficult problems related to molecular line-blanketing and the hydrodynamic structure of the atmosphere must be overcome. There is now a very large literature concerning LTE, plane-parallel, model stellar atmospheres, which cannot be described fully here;



we shall merely give a few typical references and invite the reader to examine these papers and the references cited therein. A comprehensive list, through 1965, can be found in (**506**); many of the models in that list use a grey temperature distribution on a mean optical-depth scale. Extensive grids of unblanketed, nongrey, radiative-equilibrium models can be found in (**421**; **608**); models including hydrogen-line blanketing by the "direct approach" have been calculated for A- and B-type main-sequence stars and giants (**423**; **357**), and for white dwarfs (**620**; **412**). Models for O- and B-stars, allowing for blanketing by hydrogen lines and strong lines of abundant light ions, by the direct approach, are given in (**449**; 7; **298**; **105**; **471**). Major improvements in the simulation of real atmospheres have been achieved by including the blanketing of thousands to millions of lines, using various types of opacity distribution functions. A preliminary model of Procyon (F5IV) allowed for about 30,000 lines (**612**); extensive grids including hundreds of thousands of lines semiempirically have been published in (**247**; **512**; **513**; **514**); and recently these efforts have culminated in the publication of models (**331**; **516**, 271) allowing for 1,760,000 lines on the range

$$8000°\text{K} \leqslant T_{\text{eff}} \leqslant 50{,}000°\text{K}, \quad 2 \leqslant \log g \leqslant 5$$

(as well as a solar model). A few illustrative results from these calculations will be described below.

### EMERGENT ENERGY DISTRIBUTION

The ultimate goal of stellar atmospheres analyses is the construction of mathematical models that describe the physical properties of the outer layers of stars. Having computed detailed models on the basis of the theoretical principles described in this chapter, one then compares predicted and observed values for the distribution of radiation within the spectrum, and attempts to associate a real star with a definite model. In this way values of the parameters that describe the model, ($T_{\text{eff}}$, log $g$, chemical composition), can be assigned to the star. We shall concentrate here on the comparison of observed and computed values of *continuum* features, deferring a discussion of lines to the second half of this book. In early-type stars, spectroscopic information about gravities comes mainly from profiles of the hydrogen lines (for which the broadening mechanisms are density-sensitive) and about abundances from an analysis of line-strengths; we shall therefore focus mainly on the determination of $T_{\text{eff}}$ and related parameters—e.g., the bolometric correction. In fitting the continuum we may follow several approaches.

(a) A fit can be made to the entire spectrum. This assumes that a *complete energy distribution* (perhaps including spectral regions inaccessible to ground-based observations) is available. In most cases the comparison is based on the *relative* distribution of energy in the spectrum—i.e., $F_\nu/F_{\nu_0}$, where $\nu_0$

denotes some prechosen frequency. In a few cases it is possible to make the comparison in *absolute* energy units using fluxes in ergs $cm^{-2}$ $sec^{-1}$ $hz^{-1}$ for both the star and model; here we obtain an enormously important check on the validity of the whole theory.

(b) More limited information concerning a few outstanding features in the flux distribution may be used. For example, in A- through O-stars the slope of the *Paschen continuum* (3650 Å $\leqslant \lambda \leqslant$ 8205 Å) is useful; the name is derived from the fact that the dominant opacity source on this wavelength range in early-type stars is from photoionizations of the $n = 3$ level of hydrogen. Two other important features are the *Balmer jump*,

$$D_B \equiv 2.5 \log[F_\nu(\lambda 3650^+)/F_\nu(\lambda 3650^-)]$$

and the *Paschen jump*, $D_P \equiv 2.5 \log[F_\nu(\lambda 8205^+)/F_\nu(\lambda 8205^-)]$. These parameters give measures of the effects of the onset of photoionization edges near the wavelengths stipulated.

In particular, at the Balmer jump, towards shorter wavelengths the opacity is large, owing to photoionizations from the $n = 2$ level of hydrogen, hence we receive radiation only from the upper, cooler layers; whereas towards longer wavelengths, the material is much more transparent and we see deeper, hotter, layers from which the flux is larger. The result is a fairly abrupt drop in the flux across these frequency boundaries (actually the drop is not sharp because of the opacity of overlapping lines of the series converging on the continuum). The continuum slope can be observed and computed unambiguously, but one must be able to correct the observed values for interstellar reddening effects, and must have a reliable absolute energy distribution standard (see below). The "jumps" are not as strongly affected by reddening or calibration problems because they are defined over a very limited frequency range. However, although the flux ratio is obtained easily from unblanketed models, this abstract quantity is not actually measureable, owing to the confluence of lines near the series limit; hence one must use blanketed models, and apply the *same* operational process to both observed and computed distributions to obtain meaningful comparisons.

(c) Finally, we may employ *colors* measured with filters that isolate specified bands within the spectrum. Colors can be obtained easily and accurately by standarized observational techniques, and such measures can be extended to very faint stars by use of broad-band filters. On the other hand, it is easier to calibrate colors against theoretical models for narrow bandwidths, for then one can account more accurately for line-blanketing effects in the model. In practice a compromise must be struck, and a large number of color systems with various properties exist, many of them measuring parameters that are specially designed to characterize the properties of particular groups of stars [see, e.g., the systems described in (**516**)]. A widely used system that has been well-calibrated in terms of models is the Strömgren *uvby* system.



All comparisons between models and observations rest, in the end, on the fundamental calibration of the energy distribution of a standard star (or stars) in the sky, and it is impossible to overemphasize the importance of this basic connection between theory and observation [see also (**516**, 241)]. Because it is, in practice, impossible to make an a priori determination of the absolute efficiency of the telescope–spectrometer–receiver system, one proceeds by comparing a star to a standard blackbody source of known emissivity, using the same observational apparatus. It would take us too far afield to describe the details of this procedure here; it is worth the reader's effort to study the literature on the subject [e.g., (**261**, Chap. 2; **484**; **485**; **486**; **285**; **487**; **286**; **287**; **288**) and references cited therein]! For main-sequence B-stars, both the slope of the Paschen continuum and the Balmer jump depend strongly on $T_{\text{eff}}$ and are insensitive to surface gravity (see Figure 7-3); hence both may be used to infer $T_{\text{eff}}$.

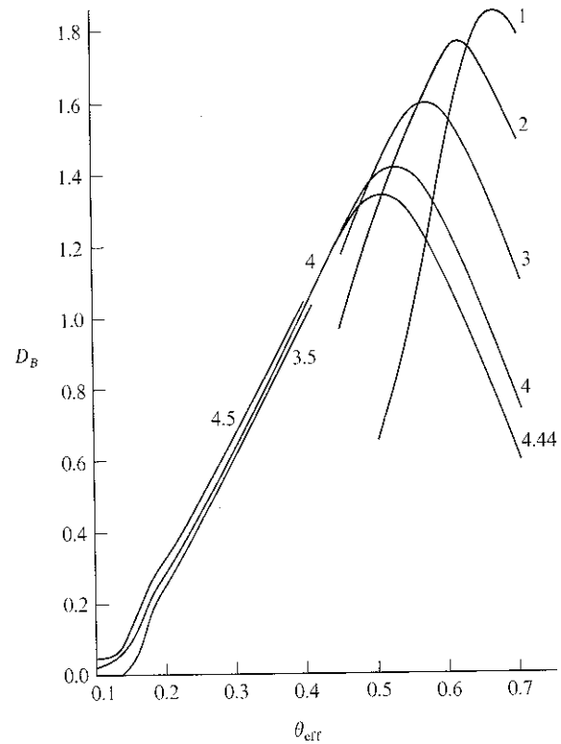



FIGURE 7-3
Balmer jumps computed from LTE model atmospheres, as a function of effective temperature and gravity. *Ordinate*: Balmer jump in magnitudes; *abscissa*: $\theta_{\text{eff}} \equiv 5040/T_{\text{eff}}$. Curves are labeled with log $g$.

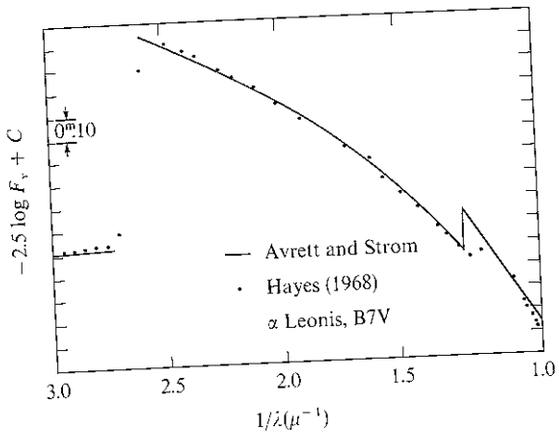



FIGURE 7-4
Comparison of the energy distribution of α Leo (B7V) as observed by Hayes (**285**) with the model atmosphere (**608**) that best fits the Paschen continuum, namely, with $T_{\text{eff}} = 13{,}000°\text{K}$ and $\log g = 4$. Note that the computed and observed Balmer jumps are consistent. *Ordinate*: relative flux in magnitudes; *abscissa*: $1/\lambda$ where $\lambda$ is in microns. From (**285**).

Until about 1968, a serious discrepancy existed between these two determinations in the sense that, if a fit was made to the Paschen continuum, the observed Balmer jump was smaller than computed (or if the Balmer jump was fitted, the slope of the observed Paschen continuum was too shallow); the discrepancy in $T_{\text{eff}}$ amounted to 3000°K (the Balmer jump temperatures being higher). The problem was resolved when a new calibration was made at Lick Observatory by Hayes (**285**; **286**), who showed that the original calibration had too flat a Paschen continuum slope. With his calibration it became possible to fit the observed spectrum very well (see Figure 7-4), and effective temperatures deduced from the two parameters were consistent [see, e.g., Figure 3 of (**682**)]; by this procedure, an effective-temperature scale for the B-stars can be established (**682**; **555**). A second recalibration made at Palomar (**487**) by Oke and Schild disagreed with the Hayes calibration (and the models) below the Balmer jump; recent work by Hayes and Latham (**287**; **288**), however, has shown conclusively that the source of the discrepancy was a faulty correction for the effects of atmospheric extinction in the Palomar data, and when this is removed, the Lick and Palomar results agree. A comparison of the energy distribution of Vega with a line-blanketed model is shown in Figure 7-5. In fitting relative energy distributions in spectral regions visible from the ground, it is important to allow for line-blanketing in spectral types later than A. For example, in Figure 7-6 we see that the blanketed model of Procyon (**612**) mentioned at the beginning of this section fits the observed



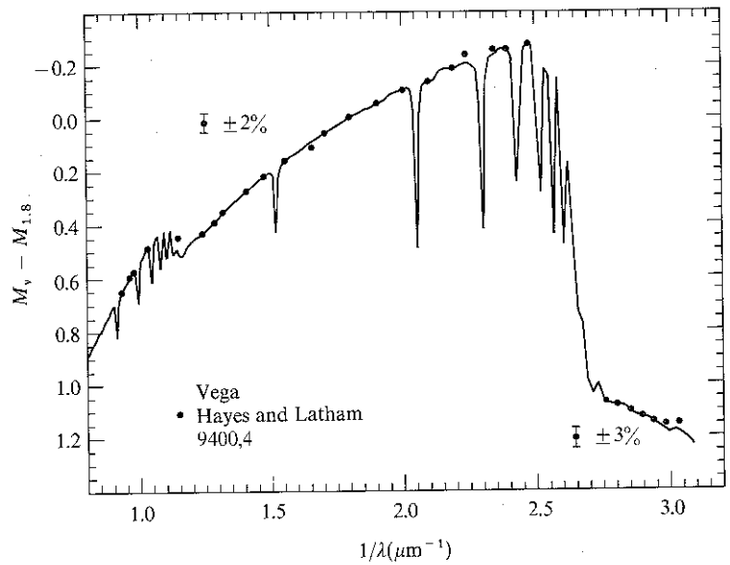



FIGURE 7-5
Comparison of energy distribution of the fundamental standard Vega as measured by Hayes and Latham (**287**) (dots) with a line-blanketed model [(**381**; **561**, 271)] that has $T_{\text{eff}} = 9400°\text{K}$ and $\log g = 4$. *Ordinate*: relative flux in magnitudes; *abscissa*: $1/\lambda$ where $\lambda$ is in microns.

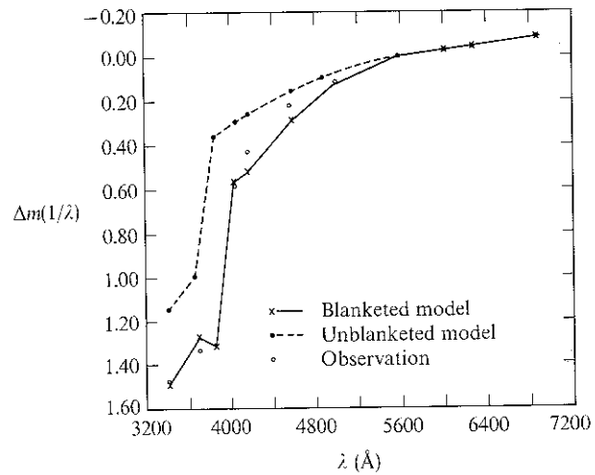



FIGURE 7-6
Unblanketed and blanketed energy distributions for models with $T_{\text{eff}} = 6500°\text{K}$ and $\log g = 4$, compared with observations of Procyon. *Ordinate*: relative flux in magnitude units; *abscissa*: wavelength $\lambda$ in Å. From (**612**), by permission.

energy distribution quite well, whereas the unblanketed model is much too bright. Blanketing effects are minor in the visible for B- and O-type stars but become large in the ultraviolet; fits made to models ignoring u.v. line-blanketing will be systematically in error (see below).

An entirely different approach to the derivation of effective temperatures can be made using *absolute* fluxes. From the fundamental calibration one can determine the actual energy output of a star at a particular wavelength; specifically, for Vega (α Lyr), which is *the* standard star, the average of the Palomar and Mt. Hopkins work yields (**287**) a flux, at the earth, of $f_\nu = 3.50 \times 10^{-20}$ ergs cm$^{-2}$ sec$^{-1}$ hz$^{-1}$ at $\lambda$5556 Å. For any other star we use the magnitude difference $\Delta m$ of the star relative to Vega (at this wavelength) to scale the flux quoted above by $10^{-0.4\Delta m}$. As was discussed in §1-4, we can convert fluxes measured at the earth to fluxes at the stellar surface if we know the angular diameter of the star. Angular diameters have been measured (**113**) for 32 stars on the spectral-type range O5 to F8; these may be used to construct an effective temperature scale. One could, for example, deduce the absolute stellar flux at some particular wavelength, and choose the model that yields the same flux to assign $T_{\text{eff}}$. By comparing the total energy emission with that observable in the visible, one can then obtain the *bolometric correction.*

Such an approach is vulnerable, however, to serious systematic errors if inadequate allowance is made for line-blanketing (**190**), and will tend to assign too-high values of $T_{\text{eff}}$ and bolometric corrections. The nature of the ultraviolet line-blanketing is illustrated in Figure 7-7. The blanketed model (**449**) allows for the strong lines of H, He, C, N, O, Si, Cl, Fe, etc. on the range 912 Å $\leqslant \lambda \leqslant$ 1600 Å by the "direct" approach. The effects of the blanketing are quite dramatic. The integrated flux of the blanketed model corresponds to $T_{\text{eff}} = 21{,}900°$K, but so much flux has been removed from the ultraviolet, and redistributed to longer wavelengths, that the energy distribution there most closely resembles an unblanketed model with $T_{\text{eff}} = 24{,}000°$K. Had we used unblanketed models to fit the visible energy distribution (whether absolute flux values or the Paschen continuum slope), the derived effective temperature would have been systematically too high by 2100°K! In fact, "direct-approach" models provide, at best, a lower bound on the total amount of blanketing, and only the recent calculations (**381**) allowing for millions of lines with opacity distribution functions provide reliable estimates of these effects.

In the face of these difficulties it is preferable to avoid direct reference to the models, and use known angular diameters, visible energy distributions, and recent space observations in the ultraviolet to construct complete absolute energy distributions *empirically* (**516**, 221; **169**). In this procedure there are nontrivial problems of ultraviolet calibrations and interstellar reddening effects but, with care, these can be overcome (**96**). From the integrated flux,

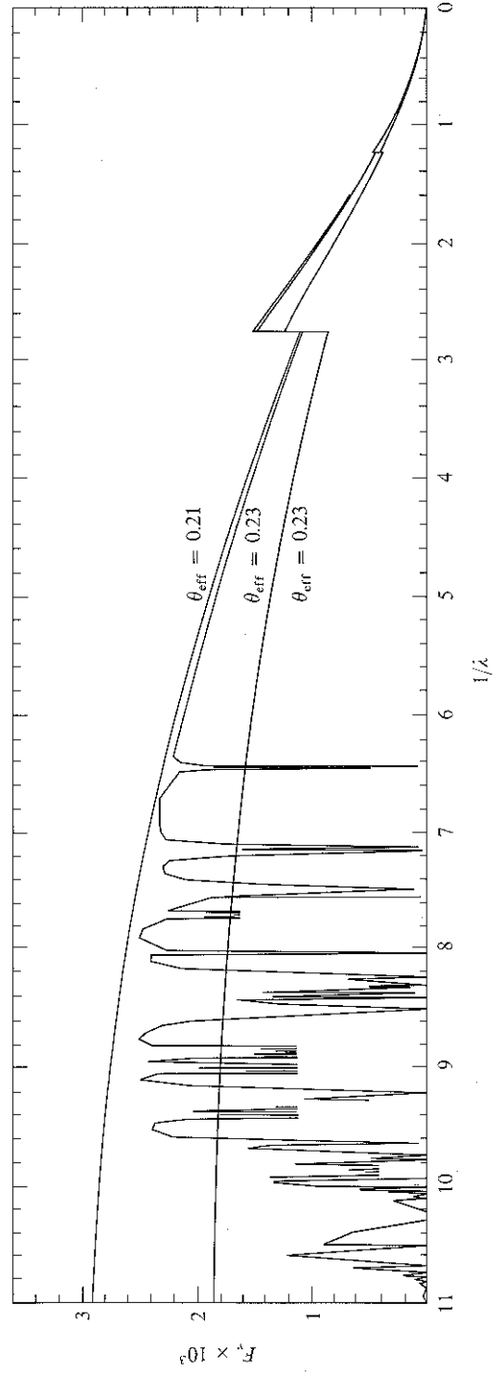



FIGURE 7-7
Flux from blanketed and unblanketed LTE models (**449**). The blanketed model yields a total energy output corresponding to $T_{\text{eff}} = 21{,}900°$K ($\theta_{\text{eff}} = 0.23$), but because the ultraviolet flux has been redistributed into the visible, the flux there is much higher than that of an unblanketed model of the same temperature, and, in fact, matches an unblanketed model with $T_{\text{eff}} = 24{,}000°$K ($\theta_{\text{eff}} = 0.21$). *Abscissa*: $1/\lambda$, where $\lambda$ is in microns; *ordinate*: $F_\nu \times 10^3$ ergs cm$^{-2}$ sec$^{-1}$ hz$^{-1}$. From (**449**), by permission.

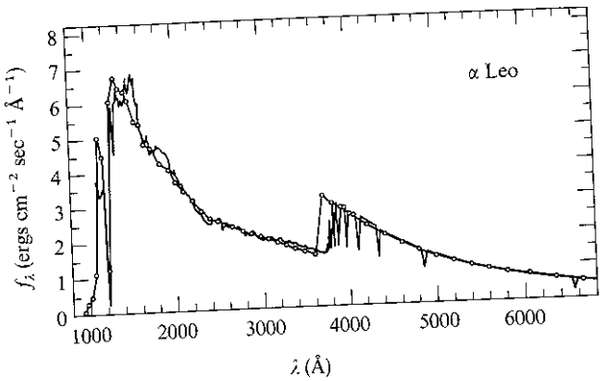



FIGURE 7-8
Comparison of empirical absolute energy distribution of α Leo (B7V) (**516**, 221) with a line-blanketed model (**381**) of the same effective temperature (12,200°K). The agreement is excellent and lends strong support to the model techniques. *Ordinate*: absolute flux $10^9$ $f_\lambda$ in ergs cm$^{-2}$ sec$^{-1}$ Å$^{-1}$ at the earth; *abscissa*: wavelength λ in Å.

the actual effective temperature is obtained; this value is essentially *independent* of any model atmosphere. A comparison of the empirical absolute energy distribution with that from a model that has the same (i.e., the empirical) $T_{\text{eff}}$ is therefore extremely significant, for it provides a test of both the absolute and relative flux predictions of the model. Such a confrontation is shown in Figure 7-8 for the B7V star α Leo (**516**, 221) and a blanketed model (**381**) of the same effective temperature. The agreement is excellent, and lends strong support to the validity of the new models.

As an example of an extreme case of line-blanketing effects, it is interesting to consider the ultraviolet flux distribution in the Ap stars as observed by OAO-2. The Ap stars are objects with anomalous abundances of certain elements (e.g., Si, Mn, Cr, Eu, Sr) that are enhanced by factors of $10^2$ to $10^3$. These stars have strong magnetic fields and show spectral variations with time; the observed variation of the field is well explained by an *oblique rotator* model in which the magnetic axis is inclined to the rotation axis of the star, while the spectral variations indicate concentrations of the elements into definite zones or patches on the stellar surface [see, e.g., (**522**; **125**; **194**)]. The greatly enhanced heavy-element abundances produce strong additional blanketing in the ultraviolet, over and above that found in normal stars. The effect is nicely illustrated in the peculiar (Si 3995) star θ Aur (see Figure 7-9), whose energy distribution in the visible matches a normal star of the same color, but in the ultraviolet (**391**) fits that of a cooler star. The effect of enhancing the line opacities in models is shown in Figure 7-10, which reproduces the behavior seen in Figure 7-9 at least semiquantitatively. Note that

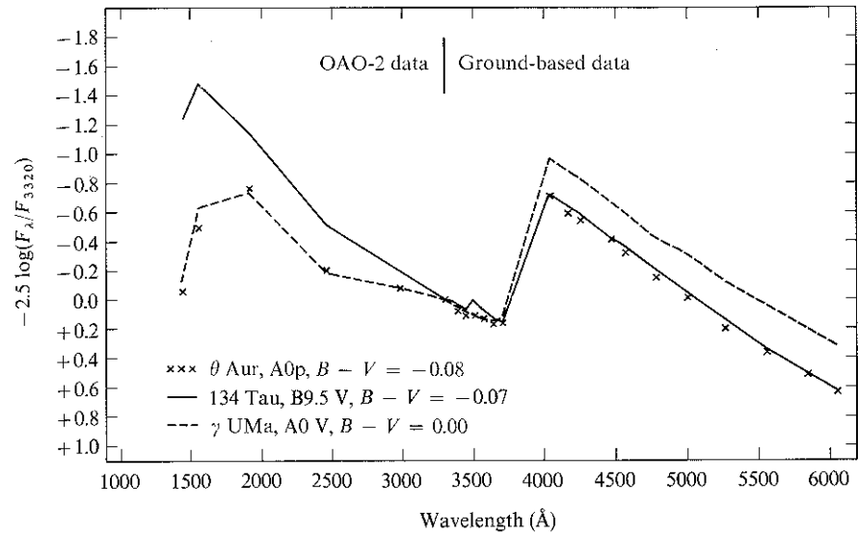



FIGURE 7-9
Comparison of relative energy distribution of the peculiar (Si 3995) A-star θ Aur with those of 134 Tau (B9.5V) and γ U Ma (A0V) (**391**). Because of the enhanced line blanketing arising from the greater heavy-element abundances in the peculiar star, the energy distribution of θ Aur matches neither of the normal stars, but resembles the cooler star in the ultraviolet and the hotter star in the visible. *Ordinate*. relative flux in magnitude units; *abscissa*: wavelength in Å. From (**391**), by permission.

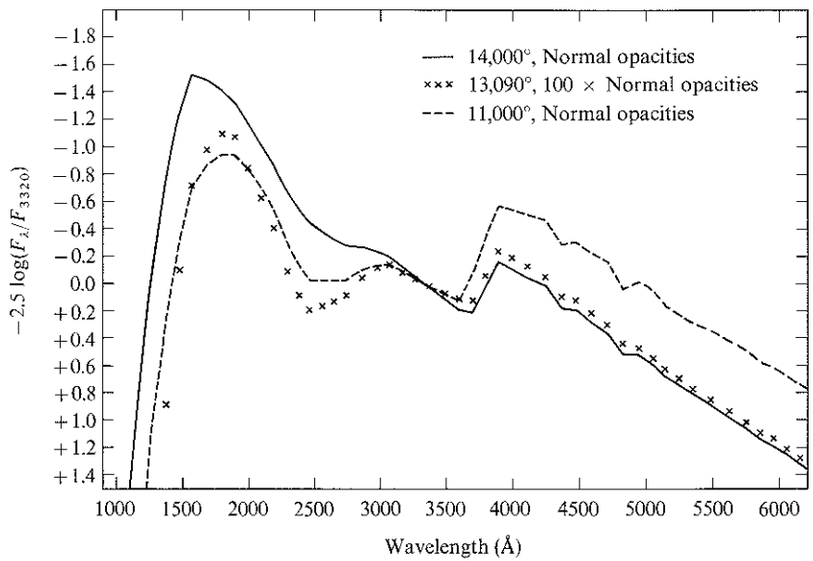



FIGURE 7-10
Line-blanketed models (**391**) showing effects of 100-fold enhanced heavy-element abundances; note strong resemblance to the effects shown in Figure 7-9. From (**391**), by permission.

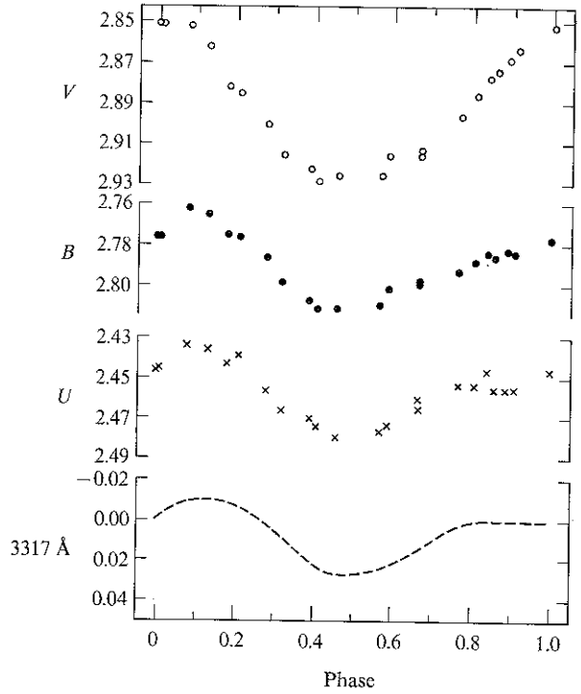



FIGURE 7-11
Variation of the peculiar (Si–Cr–Eu) A-star $\alpha^2$ CVn in UBV and at $\lambda$3317 Å as measured from OAO-2. From (**464**), by permission.

the peculiar star has a *lower* $T_{\text{eff}}$ than a normal star of the same visible color (or energy distribution), and a total energy distribution that is distinct from a normal star of the same $T_{\text{eff}}$. A further effect is shown in the Ap (Si–Cr–Eu) spectrum variable $\alpha^2$CVn. The light variations in the visible are shown in Figure 7-11 along with a near-ultraviolet band observed from OAO-2 (**464**); the far ultraviolet behavior is shown in Figure 7-12 where we see an *antiphase* variation. These results are easily interpreted in terms of much-increased ultraviolet line-blanketing at phase 0.0, which depresses the ultraviolet flux and redistributes the energy into longer wavelength bands, thus forcing a brightening in the visible; this interpretation is consistent with the fact that the rare-earth lines reach maximum strength at this phase. In contrast, at phase 0.5 we observe regions of the atmosphere where the rare-earth lines are at a minimum, hence the ultraviolet blanketing is lowest; at these phases flux emerges more freely in the ultraviolet (leading to a brightening there) and is



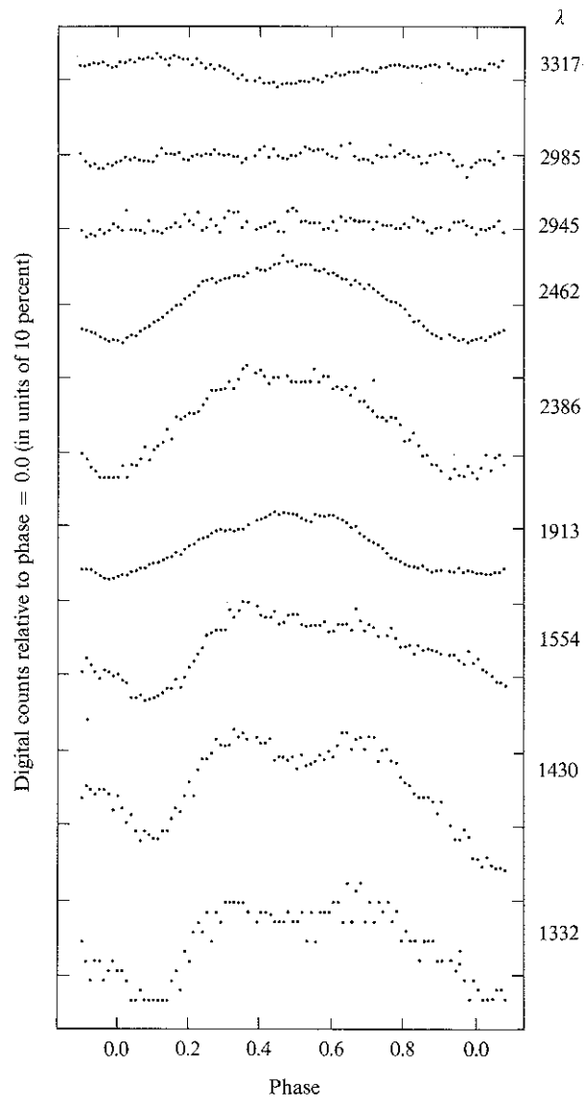



FIGURE 7-12
Variation of $\alpha^2$ CVn in ultraviolet as measured from OAO-2. Curves are labeled with wavelengths (Å) of filters. Note *antiphase* variation of far ultraviolet flux relative to visible! From (**464**), by permission.

not redistributed into the visible regions, which, accordingly, decrease in brightness. The existence of a "null wavelength" near $\lambda 2960$, which shows no variation, supports the differential line-blanketing interpretation and argues against others involving, e.g., gross geometrical deformations of the stellar surface.

For most stars we do not have detailed energy distributions, but only much more limited information such as colors measured in a photometric system. By suitable choices of filter combinations, colors can be obtained that are sensitive to effective temperature, gravity, and metal abundance and allow a determination of the amount of interstellar reddening. For example, in the Strömgren *uvby* system for, say, A–G stars, the index $(b - y)$ is a good temperature indicator, $c_1 \equiv (u - v) - (v - b)$ is gravity-sensitive, while $m_1 \equiv (v - b) - (b - y)$ is sensitive to metal abundance. To recover the information available in these data the system must be calibrated against model atmospheres. A first step in the procedure is the determination of normalizations between observed colors and those computed from the known filter transmissions. If $T_i(\lambda)$ denotes the filter transmission in color $i$, then

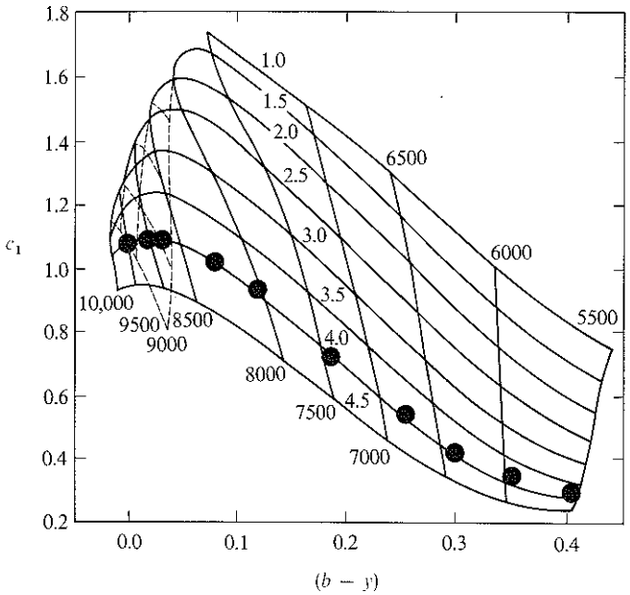



FIGURE 7-13
Comparison of observed (dots) Strömgren $(c_1, b - y)$ indices for main-sequence stars (which have $\log g \approx 4$) [(**516**, 17; **516**, 45)] with calculated values from line-blanketed models [(**381**; **516**, 271)]; note good agreement in log $g$, which suggests that $T_{\text{eff}}$ can be estimated accurately.



we have

$$(i - j)_{\text{obs}} = -2.5 \log \left[ \int_0^\infty T_i(\lambda) F(\lambda)\, d\lambda \Big/ \int_0^\infty T_j(\lambda) F(\lambda)\, d\lambda \right] + k_{ij} \quad (7\text{-}82)$$

where the constant $k_{ij}$ accounts for the unknown telescope–photometer transmission and photomultiplier response. To determine $k_{ij}$, the standard approach is to use *observed* energy distributions $F(\lambda)$ of real stars, with known colors $(i - j)_{\text{obs}}$, in equation (7-82), and to require agreement of the computed and observed colors [see, e.g., (**411**; **488**; **516**, 31; **516**, 45)]. The second step is to apply equation (7-82), with known values of $k_{ij}$, to model-atmosphere flux distributions to obtain the "observed" colors of the models. Comparisons between stellar and model colors then allow the estimation of stellar parameters; see Figure 7-13.

Once again line-blanketing plays an important role, (a) because of blocking effects in individual filter bands, and (b) because the value of $T_{\text{eff}}$ of the model depends on the effects of line-blanketing. In later-type stars it is often necessary to perform very detailed spectral synthesis studies [see, e.g., (**80**; **81**; **82**; **83**; **516**, 319)] to evaluate the blocking effects, while models with very complete opacity distribution functions such as in (**381**) are indispensible for estimating $T_{\text{eff}}$. At present considerable effort is being devoted to the development of opacity distribution functions for molecular line-blanketing; when these become available, a reliable analysis of the energy distribution of late-type stars should become possible.

### TEMPERATURE STRUCTURE

In addition to emergent fluxes, model atmospheres give the variation of the physical properties of the atmosphere with depth. In particular, we obtain an estimate of the temperature structure of the atmosphere, which, as mentioned in §7-2, is of central importance in LTE models. Let us now consider how the temperature structure in a nongrey atmosphere differs from the grey-body distribution determined in Chapter 3. We shall focus attention on two features: (a) the ratio of boundary temperature to effective temperature, $T_0/T_{\text{eff}}$, which has a value 0.811 for a grey atmosphere, and (b) the effects of backwarming. To gain insight into the physics of the situation, we shall consider two idealized problems: (1) an opacity step in the continuum, and (2) the "picket-fence" model for lines. Let us first ask qualitatively what the effects of an opacity jump or strong lines might be.

If the opacity is grey, $\chi_\nu \equiv \kappa_c + \sigma$, and the emissivity is given by $\eta_\nu = \kappa_c B_\nu + \sigma J_\nu$; then the condition of radiative equilibrium is

$$\kappa_c \int_0^\infty B_\nu(T_0)\, d\nu = \kappa_c \int_0^\infty J_\nu^0\, d\nu \quad (7\text{-}83)$$

Here the scattering term has cancelled out. From the Eddington–Barbier relation, we expect that at the surface (i.e., $\tau_\nu \ll 1$), $I_\nu^0 \approx B_\nu(\tau_\nu \approx 1)$, hence $J_\nu^0 \approx \frac{1}{2}B_\nu(T_{\text{eff}})$. Substituting this result into equation (7-83) yields the usual grey-body result $T_0^4 \approx \frac{1}{2}T_{\text{eff}}^4$. Suppose now there is a large jump in opacity (e.g., at the Lyman edge) at some critical frequency, so that $\kappa = \kappa_c$ for $\nu \leqslant \nu_0$, $\kappa = \gamma\kappa_c$ for $\nu > \nu_0$. Then the atmosphere will equilibrate to some new surface temperature $T_0'$ given by

$$\kappa_c \int_0^{\nu_0} B_\nu(T_0')\,d\nu + \gamma\kappa_c \int_{\nu_0}^{\infty} B_\nu(T_0')\,d\nu = \kappa_c \int_0^{\nu_0} J_\nu\,d\nu + \gamma\kappa_c \int_{\nu_0}^{\infty} J_\nu\,d\nu \quad (7\text{-}84)$$

Assuming that for $\nu \leqslant \nu_0$, $J_\nu \approx J_\nu^0$ (i.e., neglecting backwarming), and noting that for $\nu > \nu_0$, the surface value of $J_\nu \approx \frac{1}{2}B(T_0')$, equation (7-84) can be rewritten as

$$\begin{aligned}\kappa_c \int_0^{\infty} B_\nu(T_0')\,d\nu &\approx \kappa_c \int_0^{\infty} J_\nu^0\,d\nu \\ &\quad - \kappa_c\left[(\gamma-1)\int_{\nu_0}^{\infty} B_\nu(T_0')\,d\nu + \int_{\nu_0}^{\infty} J_\nu^0\,d\nu - \gamma\int_{\nu_0}^{\infty} J_\nu\,d\nu\right] \\ &\approx \kappa_c \int_0^{\infty} J_\nu^0\,d\nu \\ &\quad - \kappa_c\left\{\frac{\gamma}{2}\int_{\nu_0}^{\infty} B_\nu(T_0')\,d\nu + \int_{\nu_0}^{\infty}\left[\frac{1}{2}B_\nu(T_{\text{eff}}) - B_\nu(T_0')\right]d\nu\right\}\end{aligned} \quad (7\text{-}85)$$

Both of the terms in the braces are positive; hence we conclude that $T_0' < T_0$, and that the amount of cooling is larger, the larger the value of $\gamma$. This result is not rigorous because we expect $J_\nu$ must, in fact, rise above $J_\nu^0$ for $\nu \leqslant \nu_0$; but we shall see below that a rigorous analysis verifies the correctness of our conclusion. Suppose we evaluate $T'(\tau_\nu)$ at some point inside the atmosphere where $\tau_\nu \gg 1$ for $\nu > \nu_0$ while $\tau_\nu < 1$ for $\nu < \nu_0$. Then the mean intensities in the square bracket of the second equality of equation (7-85) saturate to the local Planck function and the whole bracket vanishes, and $T'(\tau_\nu)$ equals $T_0$ for the grey case; i.e., *the surface temperature drops only in those layers where the opacity jump has become transparent.*

Suppose now there are spectrum lines at frequencies $\{\nu_i\}$ that add to the opacity, $\chi_\nu = \kappa_c + \sigma + \sum_i l_i\phi_\nu$, and make both a thermal and scattering contribution to the emissivity, $\eta_\nu = \kappa_c B_\nu + \sigma J_\nu + \sum_i l_i\phi_\nu[\varepsilon_i B_\nu + (1-\varepsilon_i)J_\nu]$. Then the condition of radiative equilibrium reduces to

$$\begin{aligned}\kappa_c \int_0^{\infty} B_\nu(T_0')\,d\nu &= \kappa_c \int_0^{\infty} J_\nu^0\,d\nu \\ &\quad - \left\{\sum_i l_i\varepsilon_i[B_{\nu_i}(T_0') - \bar{J}_i] + \kappa_c \sum_i \int_{\Delta_i}(J_\nu^0 - J_\nu)\,d\nu\right\}\end{aligned} \quad (7\text{-}86)$$



where $\bar{J}_i$ denotes $\int\phi_\nu J_\nu\,d\nu$ for the $i$th line, and $\Delta_i$ is the frequency band containing that line. Again both terms in the braces are positive, so $T_0'$ must be less than $T_0$. There is the additional feature that the effect of the lines depends on their thermal coupling coefficient $\varepsilon_i$. In LTE, $\varepsilon_i \equiv 1$, $\bar{J} \approx \frac{1}{2}B_\nu(T_0')$, and for $l_i \gg \kappa_c$, a large cooling term results; thus LTE line-blanketing must drastically lower the boundary temperature. If the lines merely scatter the radiation (i.e., $\varepsilon_i \to 0$), then (just as is the case for continuum scattering!) they have no effect upon the energy balance and the boundary temperature is not changed markedly. We shall see that this conclusion is also supported by the detailed analysis to which we now turn.

The qualitative results obtained above can be put on a quantitative footing by consideration of the illuminating treatment of line-blanketing offered by the *picket-fence* model proposed by Chandraskhar (**150**) and further developed by Münch (**474**). In this model we assume (a) the continuum opacity is frequency-independent, (i.e., $\kappa_\nu \equiv \kappa$); (b) the lines have square profiles of constant width and a constant opacity ratio $\beta = l/\kappa$ relative to the continuum; and (c) the lines are distributed at random uniformly throughout the spectrum, such that within a given frequency band a fraction $w_1$ contains pure continuum, and a fraction $w_2 = 1 - w_1$ contains continuum plus lines. (Alternatively, the *probability* of finding line opacity at a specified frequency is $w_2$.) A pictorial representation of the problem is given in Figure 7-14, which shows why the name "picket-fence" is appropriate. (A slightly different interpretation of $w_1$ and $w_2$ allows treatment of an opacity step; see below.) Adopting the continuum as the standard optical depth scale, we have for

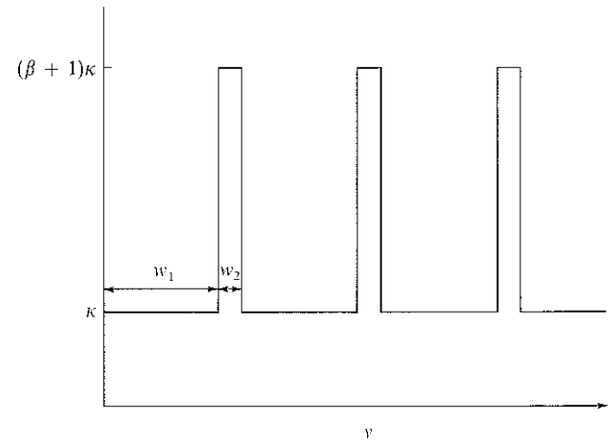



FIGURE 7-14
Picket-fence model. Lines are assumed to be a factor of $\beta$ more opaque than continuum, and to occur with a probability $w_2 = 1 - w_1$ in any frequency band.

frequencies in the continuum,

$$\mu(dI_\nu^{(1)}/d\tau) = I_\nu^{(1)} - B_\nu \tag{7-87a}$$

and, in the line

$$\mu(dI_\nu^{(2)}/d\tau) = (1 + \beta)I_\nu^{(2)} - (1 - \varepsilon)\beta J_\nu^{(2)} - (1 + \varepsilon\beta)B_\nu \tag{7-87b}$$

Integrating over all frequencies (radiation quantities without a subscript $\nu$ denote integrated quantities) and accounting for the relative probabilities that the band is covered by line or continuum, we find

$$\mu(dI^{(1)}/d\tau) = I^{(1)} - w_1 B \tag{7-88a}$$

and

$$\mu(dI^{(2)}/d\tau) = (1 + \beta)I^{(2)} - (1 - \varepsilon)\beta J^{(2)} - (1 + \varepsilon\beta)w_2 B \tag{7-88b}$$

These equations are to be solved simultaneously with a constraint of radiative equilibrium, which is obtained by integrating over angle and demanding that $F^{(1)} + F^{(2)} =$ constant, namely

$$J^{(1)} + (1 + \varepsilon\beta)J^{(2)} = [w_1 + w_2(1 + \varepsilon\beta)]B \tag{7-89}$$

Consider now the case of LTE (i.e., $\varepsilon \equiv 1$). Let $\gamma_1 \equiv 1$ and $\gamma_2 \equiv 1 + \beta$. Then equations (7-88) become

$$\mu(dI^{(l)}/d\tau) = \gamma_l(I^{(l)} - w_l B), \quad (l = 1, 2) \tag{7-90}$$

where, from equation (7-89),

$$B = \sum_{l=1}^{2} \gamma_l J^{(l)} / \sum_{l=1}^{2} w_l \gamma_l \tag{7-91}$$

To solve this system we use the discrete-ordinate approach, and choose $\{\mu_i\}$, $(i = \pm 1, \ldots, \pm n)$, such that

$$J^{(l)} = \frac{1}{2} \sum_{j=-n}^{n} a_j I_j^{(l)} \tag{7-92}$$

Then, substituting equations (7-91) and (7-92) into (7-90), we have

$$\frac{\mu_i}{\gamma_l} \frac{dI_i^{(l)}}{d\tau} = I_i^{(l)} - \left(\frac{w_l}{2\sum_m w_m \gamma_m}\right) \sum_{m=1}^{2} \gamma_m \sum_{j=-n}^{n} a_j I_j^{(m)}, \quad \begin{matrix}(l = 1, 2)\\ (i = \pm 1, \ldots, \pm n)\end{matrix} \tag{7-93}$$

If we now assume a solution of the form

$$I_i^{(l)} = C w_l e^{-k\tau}/(1 + k\mu_i/\gamma_l) \tag{7-94}$$



we find that $k$ satisfies the characteristic equation

$$\sum_{m=1}^{2} w_m \gamma_m = \sum_{m=1}^{2} w_m \gamma_m \sum_{j=1}^{n} a_j/(1 - k^2\mu_j^2/\gamma_m^2) \tag{7-95}$$

This equation yields $2n - 1$ nonzero roots for $k^2$ (bounded by poles at $1/\mu_1^2, \ldots, 1/\mu_n^2$ and $\gamma^2/\mu_1^2, \ldots, \gamma^2/\mu_n^2$) and hence $4n - 2$ values for $k$ of the form $\pm k_i$. In addition, we see by inspection that $k^2 = 0$ is a root of the characteristic equation; this root yields a particular solution of the form

$$I_i^{(l)} = b w_l(\tau + Q + \mu_i/\gamma_l) \tag{7-96}$$

which may be verified by direct substitution into equation (7-93). The general solution (7-93) is thus of the form

$$I_i^{(l)}(\tau) = w_l b\left(\tau + Q + \frac{\mu_i}{\gamma_l} + \sum_{\alpha=1}^{2n-1} \frac{L_\alpha e^{-k_\alpha\tau}}{1 + k_\alpha\mu_i/\gamma_l} + \sum_{\alpha=1}^{2n-1} \frac{L_{-\alpha} e^{k_\alpha\tau}}{1 - k_\alpha\mu_i/\gamma_l}\right), \quad \begin{matrix}(l = 1, 2)\\ (i = \pm 1, \ldots, \pm n)\end{matrix} \tag{7-97}$$

Demanding that the solution not diverge exponentially as $\tau \to \infty$, we set $L_{-\alpha} \equiv 0$ for all $\alpha$. Requiring that the total flux

$$F = 2 \sum_{l=1}^{2} \sum_{j=-n}^{n} a_j \mu_j I_j^{(l)} \tag{7-98}$$

we find

$$b = \left(\frac{3}{4} F\right) \Big/ \sum_{l=1}^{2} w_l \gamma_l^{-1} \tag{7-99}$$

The constant $Q$ and the $L_\alpha$'s are determined from the surface boundary conditions $I_{-i}^{(l)}(0) \equiv 0$, which yield a linear system of $2n$ equations in $2n$ unknowns:

$$Q - (\mu_i/\gamma_l) + \sum_{\alpha=1}^{2n-1} L_\alpha/(1 - k_\alpha\mu_i/\gamma_l) = 0, \quad \begin{matrix}(l = 1, 2)\\ (i = 1, \ldots, n)\end{matrix} \tag{7-100}$$

Using equations (7-99), (7-97), and (7-92), we find

$$J^{(l)}(\tau) = \frac{\frac{3}{4} F w_l}{\sum w_m \gamma_m^{-1}} \left(\tau + Q + \sum_{\alpha=1}^{2n-1} L_\alpha e^{-k_\alpha\tau} \sum_{j=1}^{n} \frac{a_j}{1 - k_\alpha^2\mu_j^2/\gamma_l^2}\right) \tag{7-101}$$

and, from equation (7-91),

$$B(\tau) = \frac{3}{4} F \left(\tau + Q + \sum_{\alpha=1}^{2n-1} L_\alpha e^{-k_\alpha\tau}\right) \Big/ \sum_{m=1}^{2} w_m \gamma_m^{-1} \tag{7-102}$$

We shall see below that, for the picket-fence model, $(\kappa/\bar{\kappa}_R) = \sum w_m\gamma_m^{-1}$, so from equation (7-102) we see that the asymptotic solution for $B(\tau)$ is $\frac{3}{4}F\tau_R$, as would be expected. The Rosseland mean scale $\tau_R$ *exceeds* $\tau$; in the limit of infinitely strong lines ($\gamma_2 \to \infty$), $\tau_R(\tau) = \tau/w_1$ and we see from equation (7-102) that the temperatures must be larger at depth. This is the backwarming effect, and clearly depends mainly upon the bandwidth available for continuum flux transport.

*Exercise 7-11*: (a) Verify that equation (7-96) is a particular solution of the transfer equation. (b) Verify equations (7-99), (7-101), and (7-102).

As was true for the grey problem, we may calculate the value of $B(0)$ explicitly. Define the function

$$S(x) = Q - x + \sum_{\alpha=1}^{2n-1} L_\alpha/(1 - k_\alpha x) \tag{7-103}$$

The boundary conditions (7-100) show that $S(x) = 0$ at the $2n$ values $x = \mu_i/\gamma_l$. But if we clear equation (7-103) of fractions by multiplying through by a function composed of the product of the $2n - 1$ denominators [i.e., $R(x) \equiv \prod_{\alpha=1}^{2n-1}(1 - k_\alpha x)$], then the product $R(x)S(x)$ is clearly a polynomial of order $2n$ in $x$. But we know $2n$ zeros of $S(x)$, hence the polynomial must be of the form $R(x)S(x) = C(x - \mu_1)\cdots(x - \mu_n)(x - \mu_1/\gamma)\cdots(x - \mu_n/\gamma)$. If we equate the coefficients of the two terms in $x^n$ on the two sides of this equation we can evaluate $C$ as $C = k_1k_2\cdots k_{2n-1}$, and hence obtain finally

$$S(x) = k_1\cdots k_{2n-1}\left[\prod_{i=1}^{2}\prod_{i=1}^{n}(x - \mu_i/\gamma_l)\right]\Big/\prod_{\alpha=1}^{2n-1}(1 - k_\alpha x) \tag{7-104}$$

which implies that

$$S(0) = k_1k_2\cdots k_{2n-1}\mu_1^2\mu_2^2\cdots\mu_n^2/\gamma^n \tag{7-105}$$

Now consider the characteristic function

$$\begin{aligned} T(X) &= \sum_{m=1}^{2} w_m\gamma_m\left[1 - \sum_{j=1}^{n} a_j/(1 - \mu_j^2/\gamma_m^2 X)\right] \\ &= \sum_{m=1}^{2} w_m\gamma_m\left[1 - X\sum_{j=1}^{n} a_j/(X - \mu_j^2/\gamma_m^2)\right] \\ &= \sum_{m=1}^{2} w_m\gamma_m^{-1}\sum_{j=1}^{n} a_j\mu_j^2\,[(\mu_j^2/\gamma_m^2) - X]^{-1} \end{aligned} \tag{7-106}$$

where $X \equiv 1/k^2$. We clear equation (7-106) of fractions by multiplying through by the product of the $2n$ denominators. The resulting function is a



polynomial of order $2n - 1$ in $X$; now we know that $T(X)$ has $2n - 1$ nonzero roots $X_m = 1/k_m^2$ so the polynomial must be of the form $C(X - X_1)\cdots(X - X_{2n-1})$. To evaluate $C$ we equate the coefficients of the terms in $X^{2n-1}$ to find

$$C = (-1)^{2n-1}(\sum w_m\gamma_m^{-1})(\sum a_j\mu_j^2) = (-1)^{2n-1} \times \frac{1}{3}(\sum w_m\gamma_m^{-1})$$

Thus we have

$$T(X) = \frac{1}{3}\left(\sum_{m=1}^{2} w_m\gamma_m^{-1}\right)\prod_{k=1}^{2n-1}(X_k - X)\Big/\left\{\prod_{m=1}^{2}\prod_{j=1}^{n}[(\mu_j^2/\gamma_m^2) - X]\right\} \tag{7-107}$$

From the middle equality of equation (7-106), we have $T(0) = \sum w_m\gamma_m$; and from equation (7-107), we have

$$T(0) = \frac{1}{3}(\sum w_m\gamma_m^{-1})/[(\mu_1^2\cdots\mu_n^2k_1\cdots k_{2n-1})/\gamma^n]^2$$

Combining these two results we then find, from equation (7-105),

$$S(0) = \left(\sum_{m=1}^{2} w_m\gamma_m^{-1}\right)^{\frac{1}{2}}\left(3\sum_{m=1}^{2} w_m\gamma_m\right)^{-\frac{1}{2}} \tag{7-108}$$

But comparison of equations (7-102) and (7-103) shows that

$$B(0) = \frac{3}{4}FS(0)/(\sum w_m\gamma_m^{-1}) \tag{7-109}$$

Hence we conclude that

$$[B(0)/F] = (\sqrt{3}/4)[(\sum w_m\gamma_m)(\sum w_m\gamma_m^{-1})]^{-\frac{1}{2}} \tag{7-110}$$

This result may be restated in a form that reveals its physical content. The Planck mean opacity is

$$\bar{\kappa}_P \equiv B^{-1}\int_0^\infty \kappa_\nu B_\nu\,d\nu = B^{-1}\kappa(w_1B + w_2\gamma B) = \kappa(w_1 + w_2\gamma) \tag{7-111}$$

while the Rosseland mean opacity is

$$\begin{aligned} (\bar{\kappa}_R)^{-1} &= (dB/dT)^{-1}\int_0^\infty \kappa_\nu^{-1}(dB_\nu/dT)\,d\nu \\ &= (dB/dT)^{-1}\,\kappa^{-1}(w_1 + w_2/\gamma)(dB/dT) \end{aligned} \tag{7-112}$$

or

$$\bar{\kappa}_R = \kappa(w_1 + w_2/\gamma)^{-1} \tag{7-113}$$

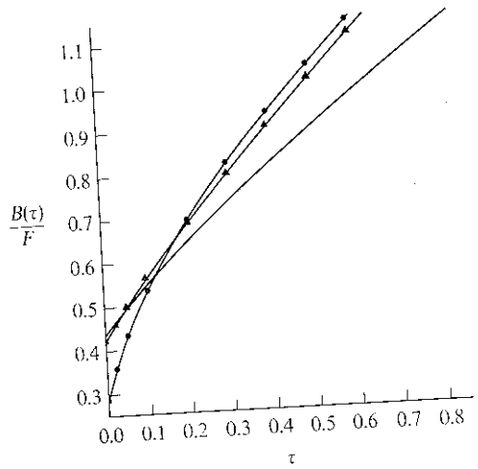



FIGURE 7-15
Depth-variation of integrated Planck function in picket-fence models. *Solid curve*: grey solution, $\beta = 1$; *solid dots*: $\gamma = 10$, $w_1 = 0.8$, $w_2 = 0.2$, $\varepsilon = 1$ (LTE); *triangles*: $\gamma = 10$, $w_1 = 0.8$, $w_2 = 0.2$, $\varepsilon = 0$ (pure scattering). Note backwarming effect in both blanketed models, the large surface-temperature drop in the LTE model, and the absence of a surface effect in the scattering model. From (**474**), by permission.

Thus equation (7-110) reduces to

$$[B(0)/F] = (\sqrt{3}/4)(\bar{\kappa}_R/\bar{\kappa}_P)^{\frac{1}{2}} \tag{7-114}$$

or

$$(T_0/T_{\text{eff}}) = (\sqrt{3}/4)^{\frac{1}{4}}(\bar{\kappa}_R/\bar{\kappa}_P)^{\frac{1}{4}} \tag{7-115}$$

Now in the limit as $\gamma \to \infty$, the Rosseland mean (being a reciprocal mean) simply saturates at a value $\kappa/w_1$ (which, in effect, shows the decrease of bandwidth available for flux transport) while the Planck mean increases without bound; thus *the effect of opaque lines in LTE is to lower the boundary temperature* (in principle to very low values). An example is shown in Figure 7-15, where $B(\tau)/F$ is plotted for the grey case and for one of Münch's solutions with $\varepsilon = 1$, $w_1 = 0.8$, $w_2 = 0.2$, and $\gamma = 10$; in this case $B(0)/F$ decreases from the grey value 0.4330 to 0.286; i.e., $T_0/T_{\text{eff}}$ drops from 0.811 to 0.721.

The analysis just described can also be applied to an opacity jump at a critical frequency, $\nu_0$, beyond which the opacity increases by a factor of $\gamma$. We now apply the two versions of equation (7-90) on the ranges ($\nu \leqslant \nu_0$) and ($\nu \geqslant \nu_0$), and define $w_1 B \equiv \int_0^{\nu_0} B_\nu \, d\nu$ and $w_2 B \equiv \int_{\nu_0}^{\infty} B_\nu \, d\nu$. We must then *assume* that $w_1$ and $w_2$ are constant with depth; for example we may



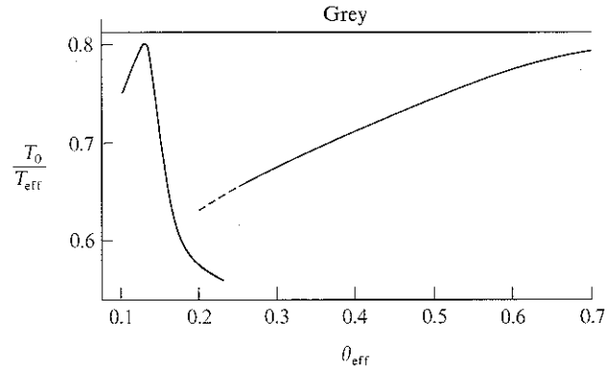



FIGURE 7-16
Ratio of boundary temperature $T_0$ to effective temperature $T_{\text{eff}}$ as a function of $\theta_{\text{eff}} \equiv 5040/T_{\text{eff}}$. The break near $\theta_{\text{eff}} = 0.25$ results from inclusion of the Lyman continuum in the high-temperature models. Upper line gives value of $T_0/T_{\text{eff}}$ for a grey atmosphere.

choose the values appropriate at $T = T_{\text{eff}}$. As pointed out by Münch (**261**, 38), this assumption is crude; but we employ it because it simplifies the analysis while retaining the essential physical content. Consider the results shown in Figure 7-16 for the ratio $T_0/T_{\text{eff}}$ from nongrey LTE model-atmospheres calculations. For all $\theta_{\text{eff}} \gtrsim 0.25$, the Lyman continuum has been omitted. For the coolest models $T_0/T_{\text{eff}}$ is near its grey value; this is not surprising because the dominant opacity source is $H^-$, which is only weakly frequency-dependent. At higher temperatures the effects of the Balmer jump become important and $T_0$ drops below its grey value. At $\theta_{\text{eff}} = 0.23$ the curve shows a sharp break caused by the effects of the Lyman continuum, which is first included at that temperature. At higher values of $T_{\text{eff}}$, hydrogen becomes more strongly ionized, the size of the Lyman jump diminishes, and the flux maximum moves beyond the jump, so $T_0/T_{\text{eff}}$ rises toward the grey value again. At still higher temperatures $T_0/T_{\text{eff}}$ drops again as a result of the He I edge at $\lambda 504$ Å and the He II edge at $\lambda 226$ Å.

We can estimate the drop in the boundary temperature caused by the Lyman jump by applying equation (7-115). Assume that bound–free and free–free absorption by hydrogen are the only sources of opacity. Using equation (4-124) for the free–free contribution, summing $n_i^* \alpha_{\nu,i}(b - f)$ over all bound levels with $u_n \equiv n^{-2}(\chi_{\text{ion}}/kT) \leqslant u \equiv (h\nu/kT)$, using equations (4-114) and (5-14), correcting for stimulated emission, and setting all Gaunt factors to unity, we may write the opacity coefficient in the form

$$\kappa_\nu^* = Cu^{-3}(1 - e^{-u})\left[1 + \sum_{u_n > u} 2u_1 n^{-3} \exp(u_1/n^2)\right] \tag{7-116}$$

where the first term in the square bracket accounts for free–free and the second for bound–free absorption. Because the Rosseland mean is a *reciprocal* mean, it will be essentially unaltered whether the Lyman continuum is included or not. Thus we need calculate only the Planck mean with and without the Lyman continuum, and use these values to estimate the ratio of $T_0$ for these two cases. We take the limits of the integral in equation (7-111) to be 0 and $u_0$, where $u_0 = u_1$ when the Lyman continuum is excluded, and $u_0 = \infty$ when it is included. Writing $B_\nu = C'u^3e^{-u}(1 - e^{-u})^{-1}$,

$$\bar{\kappa}_P(u_0) = C''\left\{1 - e^{-u_0} + \sum_n 2u_1 n^{-3}[1 - \exp(-u_0 + n^{-2}u_1)]\right\} \quad (7\text{-}117)$$

Now for $\theta_{\text{eff}} = 0.23$, $u_1 = 2.3 \times 0.23 \times 13.6 = 7.2$; if $u_0 = \infty$, the exponential term is zero identically, while if $u_0 = u_1$ it can be neglected unless $n = 1$ because $u_1 \gg 1$. Thus

$$\bar{\kappa}_P(\infty)/\bar{\kappa}_P(u_1) = \left(1 + 2u_1 \sum_{n=1}^{\infty} n^{-3}\right)\Big/\left(1 + 2u_1 \sum_{n=2}^{\infty} n^{-3}\right)$$
$$= (1 + 2.4u_1)/(1 + 0.4u_1) \quad (7\text{-}118)$$

For $u_1 = 7.2$ we thus find $\bar{\kappa}_P(\infty)/\bar{\kappa}_P(u_1) = 4.7$, hence

$$T_0(\text{Lyman cont.})/T_0(\text{no Lyman cont.}) = (4.7)^{-\frac{1}{4}} = 0.825 \quad (7\text{-}119)$$

In Figure 7-16, extrapolation of the results without Lyman continuum to $\theta_{\text{eff}} = 0.23$ yields to $T_0/T_{\text{eff}} \approx 0.65$, while $T_0/T_{\text{eff}} \approx 0.56$ when the Lyman continuum is included, which gives a boundary temperature ratio of 0.865, in good agreement with equation (7-119) (considering all the approximations that have been made). It should be noted that this temperature drop occurs only at very shallow depths where the Lyman continuum becomes transparent. At optical depths of even $10^{-4}$ in the visible, the Lyman continuum is opaque, and temperatures in models with and without the Lyman continuum are practically the same.

A further illustration of the cooling effects of LTE continua and lines is given in Figure 7-17, which shows the temperature structure in a model atmosphere with $T_{\text{eff}} = 15{,}000°\text{K}$, $\log g = 4$, consisting of hydrogen schematized as a two-level atom plus continuum (**40**). The transitions allowed in this atom are $L\alpha$, the Lyman and Balmer continua, and the free–free continuum. The temperature plateau at $T \approx 10{,}300°\text{K}$ for $-4 \leqslant \log\tau \leqslant -2$ occurs where the Balmer continuum is optically thin but the Lyman continuum is thick; this temperature gives a "$T_0$"$/T_{\text{eff}} \approx 0.68$, in fair agreement with Figure 7-16 where the Lyman continuum is omitted. Including the Lyman continuum drops $T_0$ to 9400°K; adding the $L\alpha$ line (alone) produces



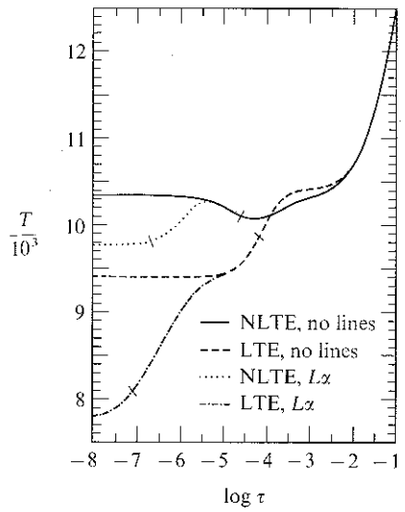



FIGURE 7-17
Temperature distribution for LTE and non-LTE models with $T_{\text{eff}} = 15{,}000°\text{K}$ and $\log g = 4$. The atmosphere is composed of hydrogen, which is represented by a schematic model atom with two bound levels and continuum. This model atom accounts for the Lyman, Balmer, and free–free continua, and the Lyman-$\alpha$ line. From (**40**), by permission.

a further drop to 7800°K. Further lines would produce yet additional cooling; the non-LTE results will be discussed in §7-5.

If we now consider *scattering* lines ($\varepsilon \neq 1$), the results obtained above change radically. Define $\lambda \equiv (1 + \varepsilon\beta)$ and $\sigma \equiv (w_1 + \lambda w_2)^{-1}$. Then equation (7-89) becomes $B = \sigma(J^{(1)} + \lambda J^{(2)})$ and equations (7-88) become

$$\mu(dI^{(1)}/d\tau) = I^{(1)} - w_1\sigma(J^{(1)} + \lambda J^{(2)}) \quad (7\text{-}88a')$$

and

$$\mu(dI^{(2)}/d\tau) = \gamma I^{(2)} - (\gamma - w_1\sigma\lambda)J^{(2)} - w_2\sigma\lambda J^{(1)} \quad (7\text{-}88b')$$

Applying the discrete-ordinate method, we obtain the characteristic equation (**474**)

$$1 - w_1\sigma G - (1 - w_1\sigma\lambda\gamma^{-1})H + w_1\sigma(1 - \lambda\gamma^{-1})GH = 0 \quad (7\text{-}120)$$

where

$$G \equiv \frac{1}{2}\sum_{i=-n}^{n} a_i/(1 + k\mu_i) \quad (7\text{-}121a)$$

and

$$H \equiv \frac{1}{2}\sum_{i=-n}^{n} a_i/(1 + k\mu_i/\gamma) \quad (7\text{-}121b)$$

Equation (7-120) has $2n - 1$ positive roots $k_\alpha$. The solution for $B(\tau)$ is

$$B(\tau) = \frac{3}{4}F\left(\tau + Q + \sum_{\alpha=1}^{2n-1} M_\alpha e^{-k_\alpha\tau}\right)\Big/\sum_m w_m\gamma_m^{-1} \quad (7\text{-}122)$$

where $M_\alpha \equiv \sigma L_\alpha[G_\alpha(1 - G_\alpha)^{-1} + (\lambda/\gamma)H_\alpha(H_\alpha - 1)^{-1}]$ and where, in turn, the $2n - 1$ constants $L_\alpha$ and the constant $Q$ are determined from the boundary conditions $I_{-i}^{(m)}(0) \equiv 0$, which imply

$$Q + w_1^{-1} \sum_{\alpha=1}^{2n-1} L_\alpha(1 - G_\alpha)^{-1}(1 - k_\alpha\mu_i)^{-1} = \mu_i, \quad (i = 1, \ldots, n) \tag{7-123a}$$

and

$$Q + (\gamma w_2)^{-1} \sum_{\alpha=1}^{2n-1} L_\alpha(H_\alpha - 1)^{-1}(1 - k_\alpha\mu_i/\gamma)^{-1} = (\mu_i/\gamma), \quad (i = 1, \ldots, n) \tag{7-123b}$$

*Exercise 7-12*: Verify equations (7-120) through (7-123).

A solution obtained by Münch for $w_1 = 0.8$, $w_2 = 0.2$, $\gamma = 10$, and $\varepsilon = 0$ is shown in Figure 7-15. Here one finds that the boundary temperature lies only slightly below its grey value, with $B(0)/F = 0.4308$ compared to the grey result 0.4330. Thus *lines, when formed by scattering, have almost no influence upon the boundary temperature*; that is, the effect of lines upon the *boundary temperature depends sensitively upon the mechanism of line-formation.* The backwarming effect is, of course, still present because the frequency band for free-flowing radiation has been restricted. In fact, the backwarming effect is almost identical in the two cases, which shows that *backwarming is determined mainly by the frequency bandwidth blocked by the lines*, and but little by details of the line-formation process. It is important to realize that LTE line-blanketing cools the surface layers (and produces darker lines); but scattering lines are also dark (cf. §10-2) even when there is no cooling at the boundary. It is *not* valid to argue for low values of $T_0$ in a stellar atmosphere just because observed lines have dark cores for, in general, the lines may be decoupled from the local temperature distribution, and their central depths may have nothing to do with $T_0$. We shall return to this point again in our work on line-formation. Finally, it is interesting to note that under certain circumstances the abrupt introduction of an opacity edge can cause local *heating* in the atmosphere [cf. (**198**)].

## 7-5 Non-LTE Radiative-Equilibrium Models for Early-Type Stars

The methods and results described thus far in this chapter have been based on the simplifying assumption of LTE. We now turn to the more general



problem of constructing models in which the populations of the atomic levels and the radiation field are computed by self-consistent solutions of the equations of transfer and of statistical equilibrium. To understand fully some of the difficulties inherent in this problem, the student should, ideally, already have mastered the material in Chapters 11 and 12; on the other hand, some of the material presented here provides background for those chapters. It is recommended, therefore, that *this section be read again after Chapters 11 and 12 are studied.* In this section we shall follow a somewhat "historical" approach in developing methods that treat, first, continuum-formation alone, and then, a final method that treats both continuum and lines. We shall not describe the line spectrum here (cf. §12-4), but will discuss the effects of lines mainly from the point of view of energy balance.

The *fundamental difficulty* in the solution of the non-LTE model-atmospheres problem is that *the occupation numbers in the outer layers of the atmosphere are determined mainly by radiative rates.* Thus the state of the material is only weakly coupled to *local* conditions (e.g., temperature and density) and is dominated by *nonlocal* information contained in the radiation field, which responds to *global* properties of the atmosphere, including boundary conditions. We shall see that the mathematical manifestation of this physical circumstance is that the source functions implied by the equations of statistical equilibrium contain *dominant* (noncoherent) *scattering terms.* We have already seen (§6-1) that these terms introduce mathematical difficulties into the solution of the transfer problem.

The first approaches to the non-LTE model atmospheres problem used an *iteration* procedure, which was successful only for continua in which the scattering terms were not large. Subsequent approaches attempted to solve the transfer equations *simultaneously* with the rate equations by introducing information from the latter explicitly into analytical expressions for the source functions used in the transfer equations. Scattering terms reduce the degree of the coupling of the material to the local thermal pool and hence also tend to introduce nonlocal information into the energy-balance criterion. Thus it becomes difficult to satisfy the requirement of radiative equilibrium. As we noted earlier in our discussion of temperature correction procedures (cf. §7-2), even small errors in energy balance may severely affect the solution of the statistical equilibrium equations. It is thus necessary to find methods that apply the *constraint* of radiative equilibrium *in addition to* the simultaneous solution of the transfer and rate equations. Initially this was done by a *linearization* procedure for the temperature alone; this procedure works when there is fairly direct coupling to the temperature structure (as there is for continua via radiative recombinations) but fails for lines where neither the emission nor absorption rates depend directly upon temperature. For models including lines it becomes necessary to make a sweeping generalization to a *complete linearization procedure* that places all physical variables

of interest on an equal footing and accounts for the global interactions of all variables throughout the atmosphere.

### SOLUTION BY ITERATION: DETAILED BALANCE IN THE LINES

The first attempts to construct non-LTE model atmospheres employed an iteration procedure in which one (a) starts with estimated occupation numbers (say from LTE, (b) uses these to compute the radiation field, and then (c) uses this radiation field to compute radiative rates in the statistical equilibrium equations, which are then solved for a new estimate of the level populations. In practice it was found that this lambda iteration procedure failed (**283**, 217) when lines were included. The lines are very weakly coupled to local conditions (see Chapter 12) and are very opaque; therefore the severe problem of *radiative* control of the populations over a very *large* range of optical depths is encountered. Just as described in §6-1 for the archetype scattering problem of the transfer equation, the iterative process then *stabilizes* to a spurious value without converging, and successive iterations differ but slightly, even though the current estimate is far from the true solution.

It is therefore of interest to inquire whether it is possible to treat *only* the continua and to ignore, or at least defer, treatment of the lines. The continua are basically simpler because (a) they are strongly coupled to local thermal conditions (via recombinations), and (b) they are relatively *transparent* down to depths where densities are high enough to assure domination by collisions (and hence recovery of LTE). This means that the self-consistency problem occurs in regions that are not optically thick, and hence that the iteration procedure has a chance of working (these remarks do *not* apply in the Lyman continuum, which is as difficult to handle as the lines). An affirmative answer to the question posed above was given by Kalkofen [(**283**, 175; **345**; **346**); see also (**424**)], who showed that, for early-type stars, the Lyman and Balmer lines are so opaque that, at depths where the visible continuum is formed, they can be expected to be in radiative detailed balance. In this case the bound–bound radiative rates upward and downward essentially cancel each other. In particular, for $T_{\text{eff}} \sim 10^4$ °K, it is found that the detailed balance criterion is met for continuum optical depths $\tau_{5000} \gtrsim 10^{-4}$, which implies that the continuum is *already formed* (i.e., is optically thin) before the lines go out of detailed balance; thus the continuum-formation problem can be treated essentially *independently* (*except* for the Lyman continuum, which is about as opaque as the lines). This result is valuable, for it offers an opportunity to assess the importance of departures from LTE from continuum observations alone.

Mathematically, radiative detailed balance implies that in the rate equations (5-87) we may *analytically cancel* out (or, equivalently, *omit*) all pairs



of terms of the form $[n_i R_{ij} - n_j(n_i/n_j)^* R_{ji}]$; we thus eliminate the most troublesome terms from the equations at the outset. Physically, the approximation proposed here recognizes that photons first "see" the surface in the most transparent regions of the spectrum (i.e., in the continuum), and that free escape out of the atmosphere in these bands leads to departures from LTE at the greatest *geometrical* depths in the star. The simplified continuum-only problem leads, therefore, to the correct asymptotic behavior at depth, and provides a starting point for the solution of problems that include the line terms.

In practice, the iteration procedure treats the departures from LTE as a perturbation away from the LTE state. Comparison of equations (7-2) and (7-4) shows that with the line-terms omitted, departures from LTE do not affect the expression for the emissivity [*if* by $n_i^*$, the LTE population of level $i$, we mean the value calculated from the Saha–Boltzmann equation (5-14) using the actual (non-LTE) electron and ion densities]. Comparison of equations (7-1) and (7-3) shows that (again omitting lines) we can write $\chi_\nu = \chi_\nu^* + \delta\chi_\nu$. If we define $b_i \equiv n_i/n_i^*$, then $\delta\chi_\nu = \sum_i d_i n_i^* \alpha_{ik}(\nu)$ where $d_i \equiv b_i - 1$. Now suppose that at any stage of the calculation we regard as given both $T(m)$ and either (a) the values of $b_i$ for all bound levels or (b) the values of all radiative continuum rates. We may then integrate the hydrostatic equation in the usual way, and solve for the electron and ion densities using either essentially the same formalism as in §5-2, but with $n_i^*$ replaced with $b_i n_i^*$ throughout, or the linearization method in §5-5 with all terms in $\delta T$ and $\delta J_k$ set to zero. The latter method yields a consistent current value for $n_e$ and $n_{\text{ion}}$; the former does not, for it ignores the nonlinearity in $n_e$ in the collision rates. In the work cited below where this iteration method was employed, the former alternative was used, and the whole process iterated to convergence.

We next solve the transfer equation

$$\mu(dI_\nu/dz) = -(\chi_\nu^* + \delta\chi_\nu)I_\nu + \kappa_\nu^* B_\nu + n_e\sigma_e J_\nu \tag{7-124}$$

or

$$\mu(dI_\nu/d\tau_\nu) = I_\nu - \xi_\nu B_\nu - \zeta_\nu J_\nu \tag{7-125}$$

where $\xi_\nu \equiv \kappa_\nu^*/(\chi_\nu^* + \delta\chi_\nu)$ and $\zeta_\nu \equiv n_e\sigma_e/(\chi_\nu^* + \delta\chi_\nu)$, using any technique that can handle the electron scattering term correctly. This yields new values for the radiation field, which are used in the rate equations to solve for new departure coefficients. For detailed balance in the lines, the complete rate equations (5-87) reduce to

$$d_i\left[4\pi\int_{\nu_0}^{\infty}(\alpha_\nu J_\nu/h\nu)\,d\nu + \sum_{j\neq i}^{\kappa} C_{ij}\right] - \sum_{j\neq i}^{I} d_j C_{ij}$$
$$= 4\pi\int_{\nu_0}^{\infty}(B_\nu - J_\nu)(1 - e^{-h\nu/kT})(\alpha_\nu/h\nu)\,d\nu \tag{7-126}$$

where $I$ denotes the last *bound* level and $d_i \equiv (b_i - 1)$. As before, we can solve these linear equations for the $d_i$'s if we regard $n_e$ and $n_{\text{ion}}$ as fixed, or we can iterate $n_e$ to consistency.

*Exercise 7-13*: Verify equation (7-126); note that, unlike equation (5-87), both upward and downward collision rates appear.

Further, temperature correction may be performed to enforce the requirement of relative equilibrium; in the papers cited below this was done using the Avrett–Krook procedure [equations (7-19) and (7-20)], modified trivially to use the generalized definitions of $\xi_\nu$ and $\zeta_\nu$ in equation (7-125) (specifically, $\zeta_\nu \neq 1 - \xi_\nu$). With the new estimate of the temperature structure and departure coefficients, the hydrostatic equilibrium and transfer equations may be solved again and the whole process iterated to convergence.

A number of models of the type described above have been constructed to study the effects of departures from LTE in the continuum and the observational implications of these departures [(**283**, 217; **348**; **610**; **425**; **426**; **427**; **452**)]. In the earliest work, very substantial changes (decreases) in the Balmer jump were predicted and it was suggested (**610**) that these effects might explain the then-existing discrepancy between observed and computed Balmer jumps (see discussion in §7-4). Subsequent work using more refined atomic models and better collision cross-sections showed that departures from LTE have negligible effects on the Balmer jump in *main-sequence* B-*stars* (important effects remaining for *supergiants* and O-*stars*), and ultimately the discrepancy was removed by a change in the fundamental calibration of the energy distribution of Vega. Nevertheless, the first papers were important, for they stimulated interest, and called attention to the possibility that departures from LTE could have observable consequences in the continuum. We defer further discussion of the observational implications to later in this section.

Results for the departure coefficients $d_i$ of the first six levels of hydrogen in a model with $T_{\text{eff}} = 10{,}000°\text{K}$, $\log g = 3$, are shown in Figure 7-18. Here we see that the deviations are, in fact, fairly small at depths representative of continuum formation. The $n = 2$ *level is underpopulated*, while levels with $n \geqslant 3$ *are overpopulated*. The values of $d_n$ decrease rapidly for large $n$ because (a) the collisional ionization rates for highly excited levels become large and force recovery of LTE, and (b) the radiation field at low frequencies is dominated by *free–free* processes—which are purely thermal, and again tend to force recovery of LTE. Then $n = 2$ level is underpopulated because, for that level at the temperatures prevalent in the model, $h\nu_2/kT \gg 1$, and the increase of temperature into the atmosphere implies that $B_\nu \sim \exp(-h\nu/kT)$ rises rapidly. Thus at the surface $J_\nu \approx \frac{1}{2}B_\nu(T_{\text{eff}})$ *exceeds* $B_\nu(T_0)$, and the level is preferentially photoionized as can be seen from equation (7-126).



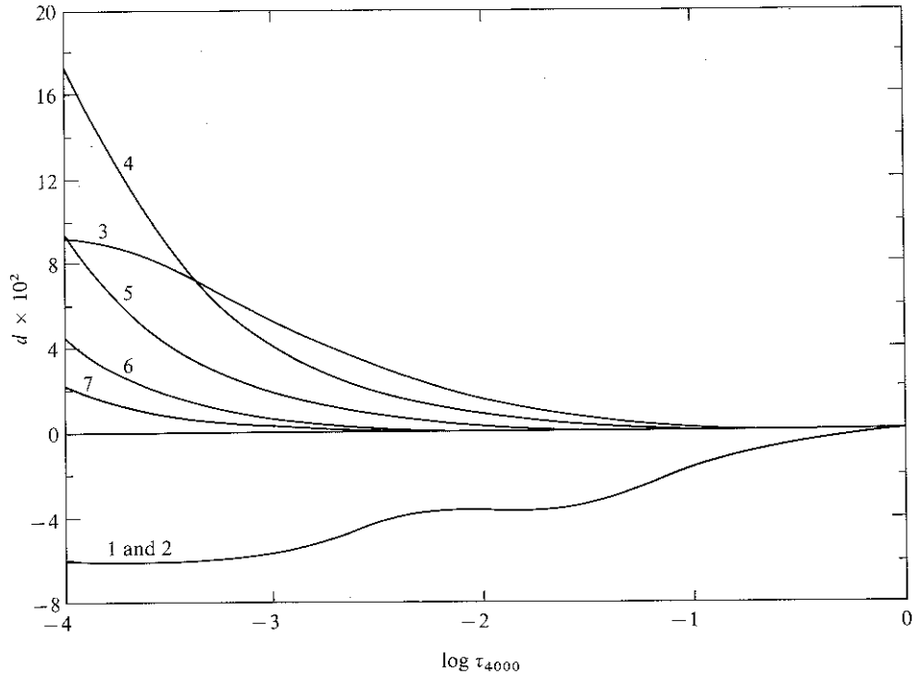



FIGURE 7-18
Non-LTE departure coefficients in the first 7 states of hydrogen for a model with $T_{\text{eff}} = 10{,}000°\text{K}$ and $\log g = 3$. Ordinate gives $d_i \times 10^2$; curves are labeled with quantum number of the level of the model atom. Note that level 2 is *underpopulated* while higher levels are *overpopulated*. Levels 1 and 2 are locked together by the assumption of detailed balance in the Lyman continuum.

On the other hand, for $n \geqslant 3$, $h\nu/kT \lesssim 1$, and the dilution factor of $\frac{1}{2}$ in $J_\nu$ outweighs the effects of the temperature gradient, so $J_\nu < B_\nu$ and the levels are overpopulated. In Figure 7-18 we have $d_1 = d_2$ because it was assumed that the Lyman continuum was also in radiative detailed balance; in this case the radiative rates in equation (7-126) for $n = 1$ cancel analytically, and we are left with $d_1 \sum_{j=2}^{\kappa} C_{1j} = \sum_{j=2}^{I} d_j C_{1j}$ which implies $d_1 \approx d_2$ because $C_{12} \gg C_{1j}$ for $j > 2$. Thus, collisional coupling of $n = 1$ to $n = 2$ allows the upper level to drive the same departure into the ground state population. Closer to the surface, where the Lyman continuum comes out of detailed balance, the $n = 1$ level becomes *overpopulated* (see below). At higher temperatures, characteristic of the O-stars (i.e., $T_{\text{eff}} \geqslant 35{,}000°\text{K}$), the situation is different, for now $n = 2$ becomes *overpopulated*, and the ground state $n = 1$ becomes *underpopulated* at depths where the Lyman continuum is formed. This would be expected on the basis of the scaling of $h\nu_0/kT_{\text{eff}}$

mentioned above; these results may also be understood (**346**) in terms of the anticipated variation of the flux with depth in the various continua (recall that $dH_\nu/d\tau_\nu = J_\nu - S_\nu$). Finally, it should be emphasized that the departure coefficients obtained from the procedure described here *cannot* be used to calculate line profiles, for they lead to spurious results, as would be expected because the level-populations will be inconsistent with the values they would have in the presence of lines (**43**).

### FORMATION OF THE LYMAN CONTINUUM

The calculations described above assume that the Lyman continuum is in radiative detailed balance. But at some point in the outer atmosphere, the Lyman continuum must begin to become transparent, and significant transfer effects occur that force $n_i R_{i\kappa}$ to depart from $n_\kappa R'_{\kappa i}$, and hence lead to an uncoupling of the $n = 1$ state from the $n = 2$ state. This situation becomes most relevant at high values of $T_{\text{eff}}$, where the high degree of ionization of hydrogen implies that the Lyman continuum is weakened to the point of being only somewhat (rather than markedly) more opaque than the visible continuum. Application of the iteration method to the Lyman continuum fails, and we can gain some important physical understanding of the problem (and also a preview of the problems of line-formation) by analyzing why this is the case. We shall see that one must account for the information in the statistical equilibrium equations by introducing them directly into the transfer equations in such a way as to yield a *simultaneous* solution of the two sets of equations.

Consider the following simplified problem. Represent the model hydrogen atom by two bound states and continuum, and assume that departures from LTE occur only in the ground state. Let the Lyman continuum threshold frequency be $\nu_0$; consider only frequencies $\nu \geqslant \nu_0$, and suppose that $h\nu_0/kT \gg 1$ so that stimulated emission can be neglected. Ignoring electron scattering and Gaunt factors, the ground-state, upper-state, and free–free opacities all have the same "profile" $\phi_\nu \equiv (\nu_0/\nu)^3$. Writing $n_1 = b_1 n_1^*$, we then have $\chi_\nu = \chi_0\phi_\nu = (b_1\chi_1^* + \chi_u^*)\phi_\nu$, where the superscript * denotes LTE values, and the subscript $u$ denotes the contribution of the upper-level and free–free continua. Similarly, $\eta_\nu = (\eta_1^* + \eta_u^*)\phi_\nu = (\chi_1^* + \chi_u^*)B_\nu\phi_\nu$. Let $d\tau_0 = -\chi_0\, dz$ be the optical depth at the continuum head. Then the transfer equation to be solved is

$$\mu(dI_\nu/d\tau_0) = \phi_\nu(I_\nu - S_\nu) \tag{7-127}$$

where
$$S_\nu = (\chi_1^* + \chi_u^*)B_\nu/(b_1\chi_1^* + \chi_u^*) = \{[(1 - r)/b_1] + r\}B_\nu \tag{7-128}$$

and
$$r \equiv \chi_u^*/(b_1\chi_1^* + \chi_u^*) \tag{7-129}$$



The ratio $r$ is generally much smaller than unity, and the essence of equation (7-128) is that the Lyman continuum source function differs from its equilibrium value by a factor of $1/b_1$. We now assume that the opacity ratio $r$ is evaluated using the *current* value of $b_1$ (as is the optical depth scale $\tau_0$), but in equation (7-128) we substitute an *analytical* expression for $(1/b_1)$ obtained from the ground-state statistical equilibrium equation:

$$(1/b_1) = \left[4\pi\int_{\nu_0}^{\infty}(\alpha_\nu/h\nu)J_\nu\, d\nu + C\right]\Big/\left[4\pi\int_{\nu_0}^{\infty}(\alpha_\nu/h\nu)B_\nu\, d\nu + C\right] \tag{7-130}$$

where $C \equiv C_{12} + C_{1\kappa}$. Substituting equation (7-130) into (7-128) shows that $S_\nu$ is of the form

$$S_\nu = \gamma_\nu\int\Phi_\nu J_\nu\, d\nu + \varepsilon_\nu B_\nu \tag{7-131}$$

That is, the source function consists of a *noncoherent scattering term* (i.e., intensities at all frequencies in the continuum are coupled) and a *thermal term* (i.e., a term independent of the radiation field). By straightforward numerical estimates [see, e.g., Table 5-1 and equation (5-44)] it is easy to show that the thermal term is *small* compared to the scattering term.

To solve equations (7-127) with (7-131) correctly, it is necessary to perform a direct solution (e.g., by the Feautrier or Rybicki methods) allowing fully for the scattering term. If equations (7-127) and (7-131) are solved directly, we have in principle obtained a *simultaneous solution* of both the transfer and the statistical equilibrium problems. Because the scattering term appears *explicitly* in the transfer equation, the correct solution is obtained (for a given run of the thermal term) over the *entire* range of optical depth in a single step, and the slow convergence properties of lambda iteration are avoided. In practice it is still necessary to iterate because the optical scale and the overlap parameter $r$ are evaluated at each stage of the calculation using current estimates of $b_1$, but experience shows that this iteration converges immediately. Calculations using an approach of this type (**426**) for B-stars show that characteristically the ground state becomes overpopulated in the outer layers (though the results in the reference cited are only schematic because the constraint of energy balance was not adequately satisfied—we shall return to this point below).

We may gain considerable insight into the problem by noting that, for $h\nu/kT \gg 1$, the frequency variation of $J_\nu$ and $B_\nu$ [roughly as $\exp(-h\nu/kT)$] shows so rapid a drop with increasing $\nu$ that practically all of the contribution to the photoionization and recombination rate integrals comes from $\nu \approx \nu_0$. Therefore, we replace the integrals by $4\pi w_0(\alpha_0/h\nu_0)J_0$ and $4\pi w_0(\alpha_0/h\nu_0)B_0$,