

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 ($\tau_v \gg 1$) in an atmosphere, and causes large departures of the mean intensity J_v from local values of the thermal source function B_v . For ease of discussion we consider a prototype source function that contains a thermal emission component and a coherent isotropic scattering term—i.e.,

$$S_v = (\kappa_v B_v + \sigma_v J_v)/(\kappa_v + \sigma_v) = (1 - \rho_v)B_v + \rho_v J_v \quad (2-39)$$

where $\rho_v \equiv \sigma_v/(\kappa_v + \sigma_v)$. The solution of the standard transfer equation

$$\mu(\partial I_v/\partial\tau_v) = I_v - S_v \quad (2-36)$$

can be written formally in terms of J_v as [cf. Exercise (2-10)]

$$J_v(\tau_v) = \Lambda_{\tau_v}[S_v] = \Lambda_{\tau_v}[B_v] + \Lambda_{\tau_v}[\rho_v(J_v - B_v)]$$

If there were *no* scattering, $\rho_v \equiv 0$, then J_v could be calculated, as a *quadrature*, from B_v ; when $\rho_v \neq 0$, we must solve an *integral equation* for J_v . One of the first methods that comes to mind to effect such a solution is *iteration*. As we know that $J_v \rightarrow B_v$ as $\tau_v \rightarrow \infty$, let us deal with $(J_v - B_v)$. Suppose ρ_v were everywhere zero; then $(J_v - B_v)$ would equal $(\bar{B}_v - B_v)$ where $\bar{B}_v(\tau_v) \equiv \Lambda_{\tau_v}[B_v]$. If ρ_v is not zero, we could regard this value as a first approximation and write

$$\begin{aligned} (J_v - B_v)^{(1)} &= (\bar{B}_v - B_v) + \Lambda_{\tau_v}[\rho_v(J_v - B_v)^0] \\ &= (\bar{B}_v - B_v) + \Lambda_{\tau_v}[\rho_v(\bar{B}_v - B_v)] \\ &= (\bar{B}_v - B_v) + \Delta^{(1)} \end{aligned} \quad (6-1)$$

Then by iteration, we find

$$(J_v - B_v)^{(n)} = (\bar{B}_v - B_v) + \sum_{i=1}^n \Delta^{(i)} \quad (6-2)$$

where $\Delta^{(n)} \equiv \Lambda_{\tau_v}[\rho_v \Delta^{(n-1)}]$. In practice we continue the iteration until some convergence criterion—e.g., $\|\Delta^{(n)}/(J_v - B_v)^{(n)}\| \leq \epsilon$, where $\epsilon \ll 1$ —is satisfied. It is clear that, if $\|\rho_v\| \ll 1$, the iteration procedure of equation (6-2) can be expected to converge, for successive corrections $\Delta^{(n)}$ must be of order $\|\rho_v\|^n$ relative to $(J_v - B_v)$. If, however, $\|\rho_v\| \approx 1$ 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 $\lambda_v \equiv 1 - \rho_v$ 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 λ_v 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 H₂ dominates the H⁻ opacity, and ρ_v is nearly unity until great depth (at some point the hydrogen rather abruptly becomes excited and ionized, and λ_v suddenly rises to unity). For lines, the corresponding thermal parameters may be very small, $\lambda_v \sim 10^{-8}$ (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 Δ 's are *monotonic*, and are *nearly equal* in iteration after iteration. In such cases, although the fractional change per iteration is ϵ ($\epsilon \ll 1$), there is no guarantee that, say, $1/\epsilon$ 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 Λ -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 Λ -operator). The failure of Λ -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

$$B_v(\tau_v) = a_v + b_v\tau_v \quad (6-3)$$

and that ρ_v is constant with depth. The zero-order moment of the transfer equation can be written, using equations (2-71) and (2-39)

$$(\partial H_v / \partial \tau_v) = J_v - S_v = \lambda_v(J_v - B_v) \quad (6-4)$$

while the first-order moment gives

$$(\partial K_v / \partial \tau_v) = H_v \quad (6-5)$$

If we use the Eddington approximation $K_v = \frac{1}{3}J_v$ and substitute equation (6-5) for H_v into equation (6-4) we obtain

$$\frac{1}{3}(\partial^2 J_v / \partial \tau_v^2) = \lambda_v(J_v - B_v) = \frac{1}{3}[\partial^2(J_v - B_v) / \partial \tau_v^2] \quad (6-6)$$

where the second equality follows from the form of B_v assumed in equation (6-3). The solution of equation (6-6) is

$$J_v - B_v = \alpha_v \exp[-(3\lambda_v)^{\frac{1}{3}}\tau_v] + \beta_v \exp[+(3\lambda_v)^{\frac{1}{3}}\tau_v] \quad (6-7)$$

As we demand that $J_v \rightarrow B_v$ as $\tau_v \rightarrow \infty$, we must have $\beta_v \equiv 0$. To evaluate α_v , we make use of the boundary condition $J_v(0) = \sqrt{3} H_v(0) = (dJ_v/d\tau_v)_0 / \sqrt{3}$ [the second equality following from equation (6-5) in the Eddington approximation]. We thus find from equation (6-7)

$$J_v(0) = a_v + \alpha_v = (dJ_v/d\tau_v)_0 / \sqrt{3} = [b_v - \alpha_v(3\lambda_v)^{\frac{1}{3}}] / \sqrt{3} \quad (6-8)$$

Hence we obtain finally

$$J_v(\tau_v) = a_v + b_v\tau_v + (b_v - \sqrt{3}a_v) \exp[-(3\lambda_v)^{\frac{1}{3}}\tau_v] / [\sqrt{3} + (3\lambda_v)^{\frac{1}{3}}] \quad (6-9)$$

Equation (6-9) reveals the essential physics of the problem. First, it shows that J_v may be markedly different from B_v at the surface. For simplicity, consider an isothermal atmosphere—i.e., $b_v = 0$ and $B_v \equiv a_v$; then at $\tau_v = 0$, $J_v(0) = \lambda_v^{\frac{1}{3}}a_v/(1 + \lambda_v^{\frac{1}{3}}) = \lambda_v^{\frac{1}{3}}B_v/(1 + \lambda_v^{\frac{1}{3}})$. Thus when $\lambda_v \ll 1$, then J_v is much smaller than B_v at the boundary. Second, we see that this departure extends deep into the atmosphere because the slow decay of the exponential term implies that $J_v(\tau_v) \rightarrow B_v(\tau_v)$ only at depths $\tau_v \gtrsim (\lambda_v)^{-\frac{1}{3}}$; in view of the small values quoted above, these are very large depths indeed. When J_v has approached B_v arbitrarily closely, we say that the solution has *thermalized*; we therefore refer to $\lambda_v^{-\frac{1}{3}}$ 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 $\lambda_v = \kappa_v/(\kappa_v + \sigma_v)$ 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 $n = 1/\lambda_v$ 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 $n^{\frac{1}{2}}\Delta\tau = \Delta\tau\lambda_v^{-\frac{1}{2}} \approx \lambda_v^{-\frac{1}{2}}$. Photons emitted at greater depths are unlikely to escape without being thermalized (hence $J_v \rightarrow B_v$), while those emitted from shallower depths manage to escape and allow J_v to fall below the thermal value (i.e., B_v).

We now can understand why Λ -iteration fails when we adopt $J_v = B_v$ as an initial estimate. Each successive iteration can propagate information about the departure of J_v from B_v 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 $\lambda_v^{-\frac{1}{3}}$ iterations to allow the effects of the boundary to make themselves felt in the solution to a thermalization depth. When

$\lambda_v \ll 1$ 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_v by a quadrature sum, and attempt to integrate numerically the system

$$\mu_i(dI_i/dt) = I_i - \frac{1}{2}\rho \sum_{j=-n}^n a_j I_j - (1 - \rho)B, \quad (i = \pm 1, \dots, \pm n) \quad (6-10)$$

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 $I_i(0) = 0$, ($i = -1, \dots, -n$), for incoming rays on the range $-1 \leq \mu_i \leq 0$ and $I_i(\tau_{\max}) = g(\mu_i)$ [e.g., $g(\mu) \equiv B$], ($i = 1, \dots, n$), for outgoing rays on the range $0 \leq \mu_i \leq 1$. Here τ_{\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 τ_{\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 τ_{\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 $\exp(\pm kt)$ where the k 's are of order $1/\mu$. In cases where the coefficients (such as ρ_v) are depth-variable, the solution no longer consists of pure exponentials, but, nevertheless, still has an exponential character, perhaps $f(\tau) \exp(\pm kt)$ where f is a weak function of τ . 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 ϵ , the parasite will be of order $\epsilon \exp(2k\tau_{\max}) \sim \epsilon 10^{kr_{\max}}$ compared to the true solution at the other boundary, and it is obvious that, unless our initial choice is very good ($\epsilon \ll 1$), 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 $n \approx kr_{\max}$ 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, τ_{\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 $\epsilon = B \exp(-2\tau_{\max})$ at the lower boundary amplifies to $\epsilon = B$ 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 $\pm\mu$:

$$\pm\mu[\partial I(z, \pm\mu, v)/\partial z] = \chi(z, v)[S(z, v) - I(z, \pm\mu, v)] \quad (6-11)$$

where we restrict μ to the half-range $0 \leq \mu \leq 1$. We now define symmetric and antisymmetric averages

$$u(z, \mu, v) \equiv \frac{1}{2} [I(z, \mu, v) + I(z, -\mu, v)] \quad (6-12)$$

and $v(z, \mu, v) \equiv \frac{1}{2} [I(z, \mu, v) - I(z, -\mu, v)] \quad (6-13)$

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

$$\mu [\partial v(z, \mu, v)/\partial z] = \chi(z, v)[S(z, v) - u(z, \mu, v)] \quad (6-14)$$

and subtracting them to obtain

$$\mu [\partial u(z, \mu, v)/\partial z] = -\chi(z, v)v(z, \mu, v) \quad (6-15)$$

Then substituting equation (6-15) into (6-14) we can eliminate v and obtain a single second-order system for u :

$$\frac{\mu^2}{\chi(z, v)} \frac{\partial}{\partial z} \left[\frac{1}{\chi(z, v)} \frac{\partial u(z, \mu, v)}{\partial z} \right] = u(z, \mu, v) - S(z, v) \quad (6-16)$$

or, defining $d\tau_v \equiv d\tau(z, v) = -\chi(z, v) dz$ and abbreviating the notation,

$$\mu^2 (\partial^2 u_{\mu\nu}/\partial \tau_v^2) = u_{\mu\nu} - S_v \quad (6-17)$$

In writing equation (6-15) we have assumed that S is symmetric in μ ; this will be true for most of the source functions we shall consider—e.g., those of the form

$$S_v = \alpha_v \int \phi_{v'} J_{v'} dv' + \beta_v \quad (6-18)$$

or $S_v = \alpha_v \int R(v', v) J_{v'} dv' + \beta_v \quad (6-19)$

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 α 's essentially stand for scattering coefficients divided by the total opacity and the β 's represent thermal terms. It must be stressed that these choices of S , 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 multi-level 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 $\tau = 0$, $I(0, -\mu, v) \equiv 0$ which implies that $v_{\mu\nu}(0) \equiv u_{\mu\nu}(0)$ so that

$$\mu (\partial u_{\mu\nu}/\partial \tau_v)_0 = u_{\mu\nu}(0) \quad (6-20)$$

At $\tau = \tau_{\max}$, we specify $I(\tau_{\max}, +\mu, v) = I_+(\mu, v)$, and write $v_{\mu\nu}(\tau_{\max}) = I_+(\mu, v) - u_{\mu\nu}(\tau_{\max})$ so that

$$\mu (\partial u_{\mu\nu}/\partial \tau_v)_{\tau_{\max}} = I_+(\mu, v) - u_{\mu\nu}(\tau_{\max}) \quad (6-21)$$

If the diffusion approximation is valid at τ_{\max} , then

$$I(\tau_{\max}, \mu, v) = B_v(\tau_{\max}) + \mu \left(\frac{1}{\chi_v} \left| \frac{\partial B_v}{\partial z} \right| \right)_{\tau_{\max}} \quad (6-22)$$

so that $u_{\mu\nu}(\tau_{\max}) = B_v(\tau_{\max})$, $v_{\mu\nu}(\tau_{\max}) = \mu (\chi_v^{-1} |\partial B_v/\partial z|)_{\tau_{\max}}$ and

$$\mu \left. \frac{\partial u_{\mu\nu}}{\partial \tau_v} \right|_{\tau_{\max}} = \mu \left(\frac{1}{\chi_v} \left| \frac{\partial B_v}{\partial z} \right| \right)_{\tau_{\max}} \quad (6-23)$$

Exercise 6-2: (a) Generalize equation (6-20) when $I(0, -\mu, v) \neq 0$. (b) Show that for a symmetric slab (infinite in x and y), of finite thickness (in z) τ_{\max} , the lower boundary condition can be written at $\tau = \frac{1}{2}\tau_{\max}$ as $(\partial u_{\mu\nu}/\partial \tau_v) = 0$. This implies that we need consider only half the slab: $0 \leq \tau \leq \frac{1}{2}\tau_{\max}$.

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 $\{\tau_d\}$, ($d = 1, \dots, D$) with $\tau_1 < \tau_2 < \dots < \tau_D$; a set of angle points $\{\mu_m\}$, ($m = 1, \dots, M$); and a set of frequency points $\{v_n\}$, ($n = 1, \dots, N$). For any variable g , we write $g(z_d, \mu_m, v_n) = g_{dmn}$. We replace integrals by

quadrature sums—e.g., for equation (6-18) we write

$$S_{dn} = \alpha_{dn} \sum_{n=1}^N a_n \phi_{dn} \sum_{m=1}^M b_m u_{dmn} + \beta_{dn} \quad (6-24a)$$

Further, we group angles and frequencies into a single serial set of values subscripted i such that $(\mu_i, v_i) = (\mu_m, v_n)$ at $i = m + (n - 1)M$, and hence reduce (6-24a) to

$$S_{di} = \alpha_{di} \sum_{l'=1}^I w_{l'} \phi_{dl'} u_{dl'} + \beta_{di} \quad (i = 1, \dots, I) \quad (6-24b)$$

Similarly, for equation (6-19) we have

$$S_{di} = \alpha_{di} \sum_{l'=1}^I \mathcal{R}_{d, l', i} u_{dl'} + \beta_{di} \quad (i = 1, \dots, I) \quad (6-25)$$

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 v (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.,

$$(dX/d\tau)_{d+\frac{1}{2}} \approx (\Delta X_{d+\frac{1}{2}}/\Delta \tau_{d+\frac{1}{2}}) = (X_{d+1} - X_d)/(\tau_{d+1} - \tau_d) \quad (6-26)$$

$$\text{and } (d^2X/d\tau^2)_d \approx [(dX/d\tau)_{d+\frac{1}{2}} - (dX/d\tau)_{d-\frac{1}{2}}] / \left[\frac{1}{2} (\Delta \tau_{d+\frac{1}{2}} + \Delta \tau_{d-\frac{1}{2}}) \right] \quad (6-27)$$

thus, defining

$$\Delta \tau_{d \pm \frac{1}{2}, i} \equiv \frac{1}{2} (\chi_{d \pm 1, i} + \chi_{di}) |z_{d \pm 1} - z_d| \quad (6-28)$$

$$\text{and } \Delta \tau_{d, i} \equiv \frac{1}{2} (\Delta \tau_{d-\frac{1}{2}, i} + \Delta \tau_{d+\frac{1}{2}, i}) \quad (6-29)$$

we rewrite equation (6-17) as

$$\begin{aligned} & \left(\frac{\mu_i^2}{\Delta \tau_{d-\frac{1}{2}, i} \Delta \tau_{d, i}} \right) u_{d-1, i} - \frac{\mu_i^2}{\Delta \tau_{d, i}} \left(\frac{1}{\Delta \tau_{d-\frac{1}{2}, i}} + \frac{1}{\Delta \tau_{d+\frac{1}{2}, i}} \right) u_{d, i} \\ & + \left(\frac{\mu_i^2}{\Delta \tau_{d, i} \Delta \tau_{d+\frac{1}{2}, i}} \right) u_{d+1, i} = u_{di} - S_{di} \quad (i = 1, \dots, I) \\ & \quad (d = 2, \dots, D - 1) \quad (6-30) \end{aligned}$$

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

$$-\mathbf{A}_d \mathbf{u}_{d-1} + \mathbf{B}_d \mathbf{u}_d - \mathbf{C}_d \mathbf{u}_{d+1} = \mathbf{L}_d \quad (6-31)$$

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

$$\mu_i(u_{2i} - u_{1i})/\Delta \tau_{\frac{1}{2}, i} = u_{1i} \quad (6-32)$$

which is only of first-order accuracy; second-order accuracy can be obtained (30) from the Taylor's expansion $u_2 = u_1 + \Delta \tau_{\frac{1}{2}}(du/d\tau)_1 + \frac{1}{2} \Delta \tau_{\frac{1}{2}}^2 (d^2u/d\tau^2)_1$, which, using equations (6-17) and (6-20), yields

$$\mu_i(u_{2i} - u_{1i})/\Delta \tau_{\frac{1}{2}, i} = u_{1i} + \left(\frac{1}{2} \Delta \tau_{\frac{1}{2}, i} / \mu_i \right) (u_{1i} - S_{1i}) \quad (6-33)$$

or, in matrix form

$$\mathbf{B}_1 \mathbf{u}_1 - \mathbf{C}_1 \mathbf{u}_2 = \mathbf{L}_1 \quad (6-34)$$

Similarly, equation (6-21) at the lower boundary becomes

$$\mu_i(u_{Di} - u_{D-1, i})/\Delta \tau_{D-\frac{1}{2}, i} = I_{Di}^+ - u_{Di} - \left(\frac{1}{2} \Delta \tau_{D-\frac{1}{2}, i} / \mu_i \right) (u_{Di} - S_{Di}) \quad (6-35)$$

which, in matrix form is

$$-\mathbf{A}_D \mathbf{u}_{D-1} + \mathbf{B}_D \mathbf{u}_D = \mathbf{L}_D \quad (6-36)$$

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

$$\begin{pmatrix} \mathbf{B}_1 - \mathbf{C}_1 \\ -\mathbf{A}_2 \quad \mathbf{B}_2 - \mathbf{C}_2 \\ 0 \quad -\mathbf{A}_3 \quad \mathbf{B}_3 - \mathbf{C}_3 \\ \vdots & \vdots & \vdots \\ -\mathbf{A}_{D-1} & -\mathbf{B}_{D-1} & -\mathbf{C}_{D-1} \\ -\mathbf{A}_D & \mathbf{B}_D & \end{pmatrix} \begin{pmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \\ \mathbf{u}_3 \\ \vdots \\ \mathbf{u}_{D-1} \\ \mathbf{u}_D \end{pmatrix} = \begin{pmatrix} \mathbf{L}_1 \\ \mathbf{L}_2 \\ \mathbf{L}_3 \\ \vdots \\ \mathbf{L}_{D-1} \\ \mathbf{L}_D \end{pmatrix} \quad (6-37)$$

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

$$\mathbf{u}_1 = \mathbf{B}_1^{-1} \mathbf{C}_1 \mathbf{u}_2 + \mathbf{B}_1^{-1} \mathbf{L}_1 \equiv \mathbf{D}_1 \mathbf{u}_2 + \mathbf{v}_1 \quad (6-38)$$

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 $\mathbf{D}_2 = (\mathbf{B}_2 - \mathbf{A}_2 \mathbf{D}_1)^{-1} \mathbf{C}_2$

$$\text{and } \mathbf{v}_2 = (\mathbf{B}_2 - \mathbf{A}_2 \mathbf{D}_1)^{-1} (\mathbf{L}_2 + \mathbf{A}_2 \mathbf{v}_1)$$

We therefore have in general

$$\mathbf{u}_d = \mathbf{D}_d \mathbf{u}_{d+1} + \mathbf{v}_d \quad (6-39)$$

$$\text{where } \mathbf{D}_d \equiv (\mathbf{B}_d - \mathbf{A}_d \mathbf{D}_{d-1})^{-1} \mathbf{C}_d \quad (6-40)$$

$$\text{and } \mathbf{v}_d \equiv (\mathbf{B}_d - \mathbf{A}_d \mathbf{D}_{d-1})^{-1} (\mathbf{L}_d + \mathbf{A}_d \mathbf{v}_{d-1}) \quad (6-41)$$

for $d = 1, \dots, D$. 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 , ($d = D - 1, \dots, 2, 1$). Having found \mathbf{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_v : 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 $1/\Delta\tau^2 \rightarrow 0$) and hence $\mathbf{J}_d \rightarrow \mathbf{S}_d$, as expected; in fact, we find $\mathbf{J} \rightarrow \mathbf{S} + \mu^2(d^2 S/dt^2)$, 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 τ ; 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 $O(n^3)$ 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_v , not $u_{\mu\nu}$, enters these constraints. We therefore eliminate the angular information by introducing variable Eddington factors $f_v \equiv K_v/J_v$ (44).

By integration of equation (6-17) over μ we obtain

$$\partial^2(f_v J_v)/\partial\tau_v^2 = J_v - S_v \quad (6-42)$$

and the boundary conditions yield

$$[\partial(f_v J_v)/\partial\tau_v]_0 = h_v J_v(0) \quad (6-43)$$

$$\text{and } \frac{\partial(f_v J_v)}{\partial\tau_v} \Big|_{\tau_{\max}} = \frac{1}{3} \left(\frac{1}{\chi_v} \left| \frac{\partial B_v}{\partial z} \right| \right)_{\tau_{\max}} \quad (6-44)$$

where $h_v \equiv H_v(0)/J_v(0)$. 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 $T_v = cDN^3$, which represents a considerable saving. To solve these equations we must know the depth-variation of f_v at all frequencies. We proceed as follows. (a) From any *given* S_v (e.g., $S_v \equiv B_v$ 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 $\mathbf{T}_v \mathbf{u}_i = \mathbf{S}_i$ where \mathbf{T} is tri-diagonal, 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 $O(n)$ operations, so the time required to evaluate the full angle-dependent radiation field for

given S_v is $T_\mu = c'DMN$. (b) Given u_{dmn} we then calculate

$$f_{dn} = \sum_m b_m \mu_m^2 u_{dmn} / \sum_m b_m u_{dmn}$$

and $h_n = \sum_m b_m \mu_m u_{1mn} / \sum_m b_m u_{1mn}$. 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_v using explicit expressions of the form of equations (6-24) and (6-25) for S_v (written in terms of J_v). We then re-evaluate S_v using the new values for J_v . (d) Because S_v 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_v 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., $S_v = \alpha J_v + \beta$. 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 $J_v \equiv 1$, 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 $J \equiv \int \phi_v J_v dv$. 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

$$\mathbf{u}_i \equiv (u_{1i}, u_{2i}, \dots, u_{Di})^T \quad (6-45)$$

where i denotes a particular angle-frequency point. Similarly let

$$\mathbf{J} \equiv (J_1, J_2, \dots, J_D)^T \quad (6-46)$$

Then at angle-frequency point i , equations (6-30), (6-33), and (6-35) yield a system

$$\mathbf{T}_i \mathbf{u}_i + \mathbf{U}_i \mathbf{J} = \mathbf{K}_i, \quad (i = 1, \dots, I) \quad (6-47)$$

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 [α_{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 \mathbf{J}_d , namely

$$\sum_{l'=1}^I w_{l'} \phi_{dl'} u_{dl'} - J_d = 0, \quad (d = 1, \dots, D) \quad (6-48)$$

Considering the grand matrix composed of all angle-frequencies and all depths, we have for the overall structure

$$\begin{pmatrix} \mathbf{T}_1 & & & & & \\ & \mathbf{T}_2 & & & & \\ & & \ddots & & & \\ & & & \mathbf{T}_I & & \\ & & & & \mathbf{U}_1 & \\ & & & & & \mathbf{u}_1 \\ & & & & & \vdots \\ & & & & & \mathbf{u}_I \\ & & & & & \mathbf{J} \\ \mathbf{V}_1 & \mathbf{V}_2 & \cdots & \mathbf{V}_I & \mathbf{E} & \end{pmatrix} = \begin{pmatrix} \mathbf{K}_1 \\ \mathbf{K}_2 \\ \vdots \\ \mathbf{K}_I \\ \mathbf{P} \end{pmatrix} \quad (6-49)$$

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

$$\mathbf{u}_i = (\mathbf{T}_i^{-1} \mathbf{K}_i) - (\mathbf{T}_i^{-1} \mathbf{U}_i) \mathbf{J} \quad (i = 1, \dots, I) \quad (6-50)$$

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 \mathbf{J} , namely $\mathbf{W}\mathbf{J} = \mathbf{Q}$ where the full $(D \times D)$ matrix \mathbf{W} is

$$\mathbf{W} \equiv \mathbf{E} - \sum_{i=1}^I \mathbf{V}_i \mathbf{T}_i^{-1} \mathbf{U}_i \quad (6-51)$$

and the vector \mathbf{Q} is

$$\mathbf{Q} = \mathbf{P} - \sum_{i=1}^I \mathbf{V}_i \mathbf{T}_i^{-1} \mathbf{K}_i \quad (6-52)$$

The final system for \mathbf{J} is solved; this yields sufficient information to calculate \mathbf{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 $O(D^2 I) = O(D^2 MN)$ operations, and solution of the final system requires $O(D^3)$ operations, so the overall computing time becomes $T_R = cD^2 MN + c'D^3$ (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., $M^3 N^3$), 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_v can be written in terms of a single quantity 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 $\mathbf{u}_i = \Lambda_i \mathbf{J} + \mathbf{M}_i$, where the Λ matrix is generated by analytical integration of the kernel function against a set of basis functions representing \mathbf{J} . In fact, \mathbf{T}_i^{-1} is the Λ_i matrix, and inversion

of \mathbf{T}_i is markedly less costly than any other approach for generating Λ (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 Λ_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 $S_v = \alpha J_v + \beta$ for the same values of v 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 \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 \mathbf{U} matrix must consist of C diagonal $(D \times D)$ matrices side by side while each \mathbf{V} matrix consists of C diagonal $(D \times D)$ matrices stacked into a column, and \mathbf{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_v is available, and hence $H_v(0) = h_v J_v(0)$ can be calculated directly. If Rybicki's method or the angle-dependent Feautrier equations are used, we can calculate $H_v(0) = \sum_m b_m \mu_m u(0, \mu_m, v)$. Alternatively, having $S_v(\tau_v)$ we can use the Φ -operator [equation (2-61)] to find $F_v(0) = \Phi_0[S(\tau_v)]$; 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 Φ_τ to S_v , or one may compute $v(\tau_{d \pm \frac{1}{2}}, \mu_m, v_n)$ 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 $\lambda 3650 \text{ \AA}$ 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 $(i \rightarrow j)$ from level i is $n_j \alpha_{ij}(v)$. Stimulated emissions return energy to the beam at a rate proportional to I_i ; 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 $n_j \alpha_{ij}(v) G(v)$ where $G(v) = g_i/g_j$ or $G(v) = (n_i/n_j)^* \exp(-hv/kT)$ 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*

$$\chi_v = \sum_i \sum_{j > i} [n_i - (g_i/g_j)n_j] \alpha_{ij}(v) + \sum_i (n_i - n_i^* e^{-hv/kT}) \alpha_{ik}(v) + \sum_k n_e n_k \alpha_{kk}(v, T) (1 - e^{-hv/kT}) + n_e \sigma_e \quad (7-1)$$

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

$$\eta_v = (2hv^3/c^2) \left[\sum_i \sum_{j > i} n_j (g_i/g_j) \alpha_{ij}(v) + \sum_i n_i^* \alpha_{ik}(v) e^{-hv/kT} + \sum_k n_e n_k \alpha_{kk}(v, T) e^{-hv/kT} \right] \quad (7-2)$$

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

$$\chi_v^* = \left[\sum_i \sum_{j>i} n_j^* \alpha_{ij}(v) + \sum_i n_i^* \alpha_{ik}(v) + \sum_k n_e n_k \alpha_{kk}(v, T) \right] \times (1 - e^{-hv/kT}) + n_e \sigma_e \quad (7-3)$$

and $\eta_v^* = (2hv^3/c^2)e^{-hv/kT}$

$$\times \left[\sum_i \sum_{j>i} n_j^* \alpha_{ij}(v) + \sum_i n_i^* \alpha_{ik}(v) + \sum_k n_e n_k \alpha_{kk}(v, T) \right] \quad (7-4)$$

Writing $\kappa_v^* = \chi_v^* - n_e \sigma_e$, we see that $\eta_v^* = \kappa_v^* B_v$, as expected from the Kirchhoff-Planck relation [equation (2-6)].

In LTE, the occupation numbers $n_i^* = n_i^*(N, T)$, hence $\chi_v^* = \chi_v^*(N, T)$ and $\eta_v^* = \eta_v^*(N, T)$, 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

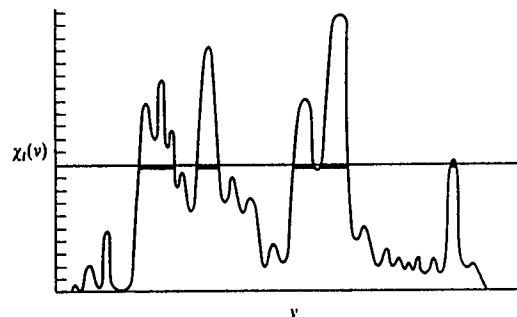


FIGURE 7-1
Schematic absorption coefficient of overlapping spectrum lines. A very large number of frequency points would be required to describe the detailed variation of the opacity.

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 $\chi_l(v) \geq X_l$, and plot a graph of this fraction against X_l . 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.

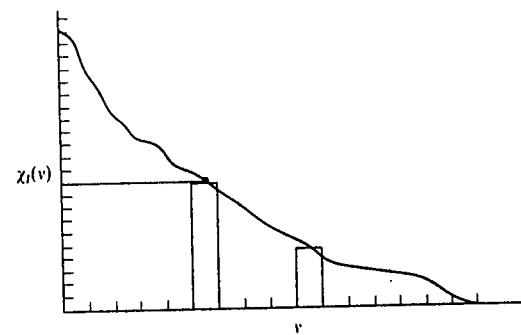


FIGURE 7-2
Schematic opacity distribution function of the spectrum in Figure 7-1. A relatively small number of representative opacities suffice to describe this smooth distribution.

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

$$\nabla p = \rho g \quad (7-5)$$

where the total pressure $p = p_g + p_R$ (dynes cm⁻²); the gas pressure $p_g \equiv NkT$; the radiation pressure $p_R \equiv (4\pi/c) \int K_v dv$; g is the surface gravity (regarded as a fundamental parameter describing the atmosphere). Here ρ is the mass density (gm cm⁻³) which, using the notation of 5-2, can be written

$$\rho = (N - n_e)m_H \sum_k \alpha_k A_k \equiv (N - n_e)\bar{m} \quad (7-6)$$

where m_H is the mass of a hydrogen atom, and A_k is the atomic weight of chemical species k with fractional abundance α_k . If we define the column mass m (gm cm⁻²) measured from the outer surface inward as our new independent variable—i.e.,

$$dm = -\rho dz \quad (7-7)$$

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:

$$(dp_g/dm) = g - (4\pi/c) \int_0^\infty (\chi_v/\rho) H_v dv \quad (7-8)$$

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_{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 $g \gtrsim 65 (T_{\text{eff}}/10^4)^4$ cm sec⁻². Clearly radiation pressure forces are negligible for the sun ($T_{\text{eff}} \approx 6 \times 10^3$, $g \approx 3 \times 10^4$) but become very important for an O-star ($T_{\text{eff}} \approx 4 \times 10^4$, $g \approx 10^4$) and in supergiants where g is quite low (and indeed approaches g_{crit}). In fact, as we shall show in Chap-

ter 15, for some O-stars the radiation forces on spectrum lines in stellar winds exceed g and accelerate the material to very large velocities (~ 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 $\sigma_e(1 + 2Y)/m_H(1 + 4Y)$. (b) Take advantage of the grey nature of electron scattering to show that, if gravity is to exceed radiation forces, then g must be $\geq g_{\text{crit}}$ where $g_{\text{crit}} = \sigma_e(1 + 2Y)(\sigma_R T_{\text{eff}}^4)/[cm_H(1 + 4Y)]$. (c) Re-express this result to show that the luminosity L of the star must be $\leq L_{\text{crit}} \approx 3.8 \times 10^4 (M/M_\odot)L_\odot$.

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

$$N_d k T_d - N_{d-1} k T_{d-1} + (4\pi/c) \sum_{n=1}^N w_n (f_{dn} J_{dn} - f_{d-1,n} J_{d-1,n}) = g(m_d - m_{d-1}) \quad (7-9)$$

Here K_v is expressed in terms of the mean intensity and a variable Eddington factor; i.e., $K_v = f_v J_v$. 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

$$N_1 k T_1 = m_1 \left[g - (4\pi/c) \sum_n w_n (\chi_{1,n}/\rho_1) h_n J_{1,n} \right] \quad (7-10)$$

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 $(\bar{\chi}/\rho)(\sigma_R T_{\text{eff}}^4/c)$ where $\bar{\chi}$ is a suitable mean opacity, then we immediately could derive the density structure $N(m)$. From this we could calculate $\chi_v^*(N, T)$, $\eta_v^*(N, T)$, 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

$$\partial^2(f_v J_v)/\partial \tau_v^2 = J_v - (\eta_v^* + n_e \sigma_e J_v)/\chi_v^* = (1 - n_e \sigma_e/\chi_v^*) J_v - (\kappa_v^*/\chi_v^*) B_v \quad (7-11)$$

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

$$4\pi \int_0^\infty [\eta_v^* - (\chi_v^* - n_e \sigma_e) J_v] dv = 4\pi \int_0^\infty \kappa_v^* (B_v - J_v) dv = 0 \quad (7-12)$$

or, in discrete form, (and allowing for departures from LTE),

$$4\pi \sum_n w_n [\eta_n - (\chi_n - n_e \sigma_e) J_n] = 0 \quad (7-13)$$

In radiative equilibrium the total flux $4\pi H = \sigma_R T_{\text{eff}}^4 = \text{constant}$, and we may choose it (or T_{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 $J_v = \Lambda_{\tau_v} [B_v(T_0)]$, 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

$$\int_0^\infty \kappa_v^* B_v(T_0 + \Delta T) dv = \int_0^\infty \kappa_v^* J_v dv \quad (7-14)$$

Expanding $B_v(T_0 + \Delta T) \approx B_v(T_0) + \Delta T(\partial B_v/\partial T)$ we find

$$\Delta T \approx \int_0^\infty \kappa_v^* [J_v - B_v(T_0)] dv / \int_0^\infty \kappa_v^* (\partial B_v/\partial T)_{T_0} dv \quad (7-15)$$

It must be emphasized that J_v in equation (7-15) denotes the value already computed from $B_v(T_0)$. 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 $J_v = \Lambda_{\tau_v} [B_v(T_0)] = B_v(T_0) + O(e^{-\tau_v})$, 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 κ_v^* 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 $\tau_v \rightarrow 1$ while the contribution to the denominator swamps that of all other bands. In effect the Λ -iteration procedure is effective only over $\Delta \tau_v \sim 1$ 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 ΔT , computed at some depth τ , has on $J_v(\tau')$ at all other depths (i.e., J_v is presumed to be fixed). This oversight necessarily leads to spurious values of ΔT . Actually $J_v(\tau_v) = \Lambda_{\tau_v} [B_v(T + \Delta T)]$, which means we should, in reality, solve an integral equation for ΔT ; 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 $d\tau \equiv -\kappa_v^* dz$, then exact frequency integrals of the moment equations (using quantities

without subscripts to denote frequency-integrated variables) are

$$(dH/d\tau) = (\kappa_f^*/\kappa_p^*)J - B \quad (7-16)$$

and

$$(dK/d\tau) = (\chi_f^*/\kappa_p^*)H \quad (7-17)$$

where κ_f^* is the absorption mean [equation (3-32)] and χ_f^* is the flux mean [equation (3-21)]; note that the scattering coefficient is included in χ_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

$$\Delta B(\tau) = -d(\Delta H)/d\tau + (\kappa_f^*/\kappa_p^*) \left[3 \int_0^\tau (\chi_f^*/\kappa_p^*) \Delta H(\tau') d\tau' + 2 \Delta H(0) \right] \quad (7-18)$$

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 A-iteration procedure; the other terms introduce new information that gives nonnegligible values of ΔB 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 τ_1 and T_1 , and by taking moments of the resulting first-order equation of the perturbation expansion, equations are derived for τ_1 and T_1 . These equations [extended to allow for scattering terms (421) and with an improved closure relation (351)] are

$$\begin{aligned} \tau'_1 + \tau_1 & \left(\int_0^\infty \chi'_v H_v^0 dv / \int_0^\infty \chi_v H_v^0 dv \right) \\ & = (1 - \mathcal{H}/H^0) - 3^{-\frac{1}{2}} \int_0^\infty \chi_v [J_v^0 - B_v(T_0)] dv / \int_0^\infty \chi_v H_v^0 dv \quad (7-19) \end{aligned}$$

and

$$\begin{aligned} T_1 = & \left\{ (1 + \tau_1) \int_0^\infty \chi_v (1 - \rho_v) [J_v^0 - B_v(T_0)] dv \right. \\ & - 3^{\frac{1}{2}} (1 - \mathcal{H}/H^0) \int_0^\infty \chi_v (1 - \rho_v) H_v^0 dv \\ & \left. + \tau'_1 \int_0^\infty [\chi'_v (1 - \rho_v) - \chi_v \rho'_v] [J_v^0 - B_v(T_0)] dv \right\} / \int_0^\infty \chi_v (1 - \rho_v) (\partial B_v / \partial T)_0 dv \quad (7-20) \end{aligned}$$

where the prime denotes the derivative d/dt ; quantities with superscript or subscript zero denote current values; $H^0 \equiv \int_0^\infty H_v^0 dv$; \mathcal{H} denotes the nominal flux ($\sigma T_{\text{eff}}^4 / 4\pi$); and $\rho_v \equiv \sigma_v / \chi_v = \sigma_v / (\kappa_v + \sigma_v)$. Equation (7-19) is a linear first-order differential equation for τ_1 that may be integrated straight-away starting with $\tau_1(0) = 0$; given values for τ_1 and τ'_1 , one then computes T_1 from equation (7-20) and hence obtains $T(\tau) = T_0(\tau) + T_1(\tau)$ at $\tau = t + \tau_1(\tau)$. Experience shows that the τ_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 $|d \ln F/dt|$ 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 $\int_{v_0}^{\infty} \alpha_v^{-1} J_v dv = \int_{v_0}^{\infty} \alpha_v^{-1} B_v dv$ while the net radiative rate in the statistical equilibrium calculation is of the form $\int_{v_0}^{\infty} \alpha_v^{-1} [b_1 J_v - B_v] v^{-1} dv$ (where b_1 denotes n_1/n_0^* of the ground state). If $b_1 \approx 1$, these two criteria differ only by the (weak!) "profile" function v^{-1} . In the limit that $h\nu_0/kT \gg 1$ and both B_v and J_v 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 ΔT ,

$$\int_0^{\infty} \kappa_v^* B_v(T + \Delta T) dv = \int_0^{\infty} \kappa_v^* \Lambda_{\tau_v} [B_v(T + \Delta T)] dv \quad (7-21)$$

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 Λ -operator; we introduce a discrete set of points $\{\tau_{v,d}\}$, ($d = 1, \dots, D$), at which we wish to determine the solution, and we represent the variation of $B_v(\tau_v)$ analytically in terms of interpolation (basis) functions on this mesh. The integral

$$J_v(\tau_v) = \frac{1}{2} \int_0^{\infty} E_1 |t_v - \tau_v| B_v(t_v) dt_v \quad (7-22)$$

is then calculated analytically for these basis functions to produce the system

$$J_{d,v} = \sum_{d'=1}^D \Lambda_{dd',v} B_{d',v} + M_{d,v}, \quad (d = 1, \dots, D) \quad (7-23)$$

where the $M_{d,v}$ represent the contribution of the interval $(\tau_{D,v}, \infty)$ to the integral. Substituting equation (7-23) into (7-21), and (a) assuming κ_v^* is unchanged by ΔT , (b) writing $B_v(T + \Delta T) \approx B_v(\tau) + (\partial B_v / \partial T) \Delta T$, and (c) introducing a frequency quadrature $\{\nu_n\}$, ($n = 1, \dots, N$), we find a set of linear equations for the values of ΔT_d :

$$\begin{aligned} & \sum_{d'=1}^D \left[\sum_{n=1}^N w_n \kappa_{d'n}^* (\partial B / \partial T)_{d'n} (\delta_{dd'} - \Lambda_{dd'n}) \right] \Delta T_{d'} \\ &= \sum_n w_n \kappa_{dn}^* \left[M_{dn} + \sum_{d'=1}^D (\Lambda_{dd'n} - \delta_{dd'}) B_{d'n} \right], \quad (d = 1, \dots, D) \end{aligned} \quad (7-24)$$

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_v is only linear, the system has to be iterated to convergence by using the new temperatures to recalculate B_v , $(\partial B_v / \partial T)$, κ_v^* , 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 Λ matrix is cumbersome and costly, and (because κ_v^* really is a function of T) must be done again for each iteration. (2) It is possible to calculate the response of the Λ matrix to changes in τ_v (arising from changes in κ_v^* 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 $\tau_v > \tau_{D,v}$

$$B_v(\tau_v) = B_v(T_D) + \frac{1}{\kappa_v^*} \frac{\partial B_v}{\partial T} \left| \frac{dT}{dz} \right| (\tau_v - \tau_{D,v}) \quad (7-25)$$

$$I(\tau_v, \mu) \approx B_v(T_D) + \frac{\mu}{\kappa_v^*} \frac{\partial B_v}{\partial T} \left| \frac{dT}{dz} \right| \quad (7-26)$$

Integrating against μ , and over all frequencies, we find

$$H = \frac{1}{3} \left(\int_0^{\infty} \frac{1}{\kappa_v^*} \frac{\partial B_v}{\partial T} dv \right) \left| \frac{dT}{dz} \right| \quad (7-27)$$

which fixes dT/dz in equation (7-25), and introduces the flux into the quantities $M_{d,v}$; 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 $B_v(\tau_v)$ 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: $\kappa_v = \kappa_0$ for $(v \leq v_0)$, $\kappa_v = \alpha\kappa_0$ for $(v > v_0)$; 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 Λ matrix and M vector are independent of frequency (though different) in the two ranges ($v \leq v_0$), ($v > v_0$); hence the integrals over frequency of B_v and $\partial B_v/\partial T$ 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 $(0, v_0)$ when $\beta \ll 1$ and for (v_0, ∞) when $\beta \gg 1$] or in terms of Debye integrals (4, 998). Solve the problem for several values of α and β , 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

$$\sum_n w_n \kappa_{dn}^* J_{dn} / \sum_n w_n \kappa_{dn}^* B_{dn} \equiv 1 \quad (7-28)$$

at all depths d , he solved the transfer problem [equation (6-30) or (6-42)] with the source function

$$S_{dn} = B_{dn} \cdot \left(\sum_{n'} w_{n'} \kappa_{dn'}^* J_{dn'} / \sum_{n'} w_{n'} \kappa_{dn'}^* B_{dn'} \right) \quad (7-29)$$

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_v and κ_v^* , equations (7-29) and the discretized form of (6-42) are solved for J_v 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_v and κ_v^* in terms of ΔT 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 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_v^* the corresponding Planck function, then the solution of the transfer equation

$$\partial^2(f_v J_v)/\partial\tau_v^2 = J_v - B_v^* \quad (7-30)$$

with lower boundary condition [using equation (7-27)]

$$\frac{\partial(f_v J_v)}{\partial\tau_v} = \left(\frac{H}{\kappa_v^*} \frac{\partial B_v^*}{\partial T} \right) / \left(\int_0^\infty \frac{1}{\kappa_v^*} \frac{\partial B_v^*}{\partial T} dv \right) \quad (7-31)$$

would automatically satisfy the condition of radiative equilibrium

$$\sum_n w_n \kappa_{dn}^* J_{dn} = \sum_n w_n \kappa_{dn}^* B_{dn}^* \quad (7-32)$$

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 κ_v^* , expand $B_v^* = B_v(T) + (\partial B_v/\partial T) \Delta T$ and $(\partial B_v^*/\partial T) = (\partial B_v/\partial T) + (\partial^2 B_v/\partial T^2) \Delta T$ in equations (7-30) and (7-31), where ΔT must satisfy

$$\Delta T_d = \sum_n w_n \kappa_{dn}^* (J_{dn} - B_{dn}) / \sum_n w_n \kappa_{dn}^* (\partial B_n/\partial T)_d \quad (7-33)$$

We now can follow two possible approaches. We could expand equations (7-30) and (7-31) as described, eliminate ΔT using equation (7-33), and solve

the system

$$\frac{\partial^2(f_n J_{dn})}{\partial \tau_n^2} \Big|_d = J_{dn} - B_{dn} - \left(\frac{\partial B_n}{\partial T} \right)_d \frac{\sum_n w_n \kappa_{dn}^*(J_{dn} - B_{dn})}{\sum_n w_n \kappa_{dn}^* (\partial B_n / \partial T)_d} \quad (7-34)$$

with the corresponding boundary condition, regarding $J_{dn'}$ as unknown.

Exercise 7-7: Write out the perturbed boundary condition with ΔT 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 ΔT , it is more efficient to use Rybicki's method, letting ΔT be the constraint variable replacing 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 ΔT , 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 $T^4(\bar{\tau}_R) = \frac{3}{4}T_{\text{eff}}^4[\bar{\tau}_R + q(\bar{\tau}_R)]$ 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

$$(dp_g/dm) = g - (\sigma_R T_{\text{eff}}^4/c)(\bar{\chi}_R/\rho) \quad (7-35)$$

simultaneously with the definition of the optical-depth scale

$$d\bar{\tau}_R = (\bar{\chi}_R/\rho) dm \quad (7-36)$$

step-by-step on a mesh $\{m_d\}$, ($d = 1, \dots, D$). 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 χ_{dn}^* and η_{dn}^* from equations (7-3) and (7-4) on a frequency mesh $\{v_n\}$, 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

$$(f_{2n} J_{2n} - f_{1n} J_{1n})/\Delta\tau_{\frac{1}{2},n} = h_n J_{1n} \quad (7-37a)$$

$$\begin{aligned} \frac{f_{d-1,n} J_{d-1,n}}{\Delta\tau_{d-\frac{1}{2},n} \Delta\tau_{dn}} - \frac{f_{dn}}{\Delta\tau_{dn}} \left(\frac{1}{\Delta\tau_{d-\frac{1}{2},n}} + \frac{1}{\Delta\tau_{d+\frac{1}{2},n}} \right) J_{dn} + \frac{f_{d+1,n} J_{d+1,n}}{\Delta\tau_{d+\frac{1}{2},n} \Delta\tau_{dn}} \\ = \left(1 - \frac{n_{e,d} \sigma_e}{\chi_{dn}} \right) J_{dn} - \frac{\eta_{dn}}{\chi_{dn}}, \quad (d = 2, \dots, D-1) \end{aligned} \quad (7-37b)$$

and (at $d = D$),

$$\begin{aligned} (f_{Dn} J_{Dn} - f_{D-1,n} J_{D-1,n})/\Delta\tau_{D-\frac{1}{2},n} \\ = H \chi_{Dn}^{-1} (\partial B_v / \partial T)_{Dn} / \sum_n w_n [\chi_{Dn}^{-1} (\partial B_v / \partial T)_{Dn}] \end{aligned} \quad (7-37c)$$

$$\text{where } \Delta\tau_{d\pm\frac{1}{2},n} \equiv \frac{1}{2} [(\chi_{d\pm\frac{1}{2},n} / \rho_{d\pm\frac{1}{2}}) + (\chi_{dn} / \rho_{dn})] |m_{d\pm\frac{1}{2}} - m_d| \quad (7-38a)$$

$$\Delta\tau_{dn} \equiv \frac{1}{2} (\Delta\tau_{d-\frac{1}{2},n} + \Delta\tau_{d+\frac{1}{2},n}) \quad (7-38b)$$

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 δT_d implies a change δN_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 δ '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 δJ 's from the constraint equations (radiative and hydrostatic equilibrium), yielding a final system for the perturbations of the "fundamental" variables δN and δT . Thus, linearizing the transfer equation (assuming the Eddington factors remain unchanged) we have, away from the boundaries, and at each frequency ν_n ,

$$\begin{aligned} \frac{f_{d-1,n} \delta J_{d-1,n}}{\Delta \tau_{d-\frac{1}{2},n} \Delta \tau_{dn}} - \left[\frac{f_{dn}}{\Delta \tau_{dn}} \left(\frac{1}{\Delta \tau_{d-\frac{1}{2},n}} + \frac{1}{\Delta \tau_{d+\frac{1}{2},n}} \right) + \left(1 - \frac{n_{e,d} \sigma_e}{\chi_{dn}} \right) \right] \delta J_{dn} \\ + \frac{f_{d+1,n} \delta J_{d+1,n}}{\Delta \tau_{d+\frac{1}{2},n} \Delta \tau_{dn}} + a_{dn} \delta \omega_{d-1,n} + b_{dn} \delta \omega_{dn} + c_{dn} \delta \omega_{d+1,n} \\ - (\eta_{dn} + n_{e,d} \sigma_e J_d) \frac{\delta \chi_{dn}}{\chi_{dn}^2} + \frac{\delta \eta_{dn}}{\chi_{dn}} + \frac{\sigma_e J_{dn}}{\chi_{dn}} \delta n_{e,d} \\ = \beta_{dn} + J_{dn} - (n_{e,d} \sigma_e J_{dn} + \eta_{dn}) / \chi_{dn} \end{aligned} \quad (7-39)$$

$$\text{where } \alpha_{dn} \equiv (f_{dn} J_{dn} - f_{d-1,n} J_{d-1,n}) / (\Delta \tau_{d-\frac{1}{2},n} \Delta \tau_{dn}) \quad (7-40)$$

$$\gamma_{dn} \equiv (f_{dn} J_{dn} - f_{d+1,n} J_{d+1,n}) / (\Delta \tau_{d+\frac{1}{2},n} \Delta \tau_{dn}) \quad (7-41)$$

$$\beta_{dn} \equiv \alpha_{dn} + \gamma_{dn} \quad (7-42)$$

$$a_{dn} \equiv \left[\alpha_{dn} + \frac{1}{2} \beta_{dn} (\Delta \tau_{d-\frac{1}{2},n} / \Delta \tau_{dn}) \right] / (\omega_{d-1,n} + \omega_{dn}) \quad (7-43)$$

$$c_{dn} \equiv \left[\gamma_{dn} + \frac{1}{2} \beta_{dn} (\Delta \tau_{d+\frac{1}{2},n} / \Delta \tau_{dn}) \right] / (\omega_{dn} + \omega_{d+1,n}) \quad (7-44)$$

$$b_{dn} \equiv a_{dn} + c_{dn} \quad (7-45)$$

$$\text{and } \omega_{dn} \equiv \chi_{dn} / \rho_d \quad (7-46)$$

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 δN and δT . Thus, from equations (7-6) and (7-3)

$$\delta \rho_d = \bar{m} (\delta N_d - \delta n_{e,d}) \quad (7-47)$$

$$\delta \chi_{dn}^* = (\partial \chi_n^* / \partial T)_d \delta T_d + (\partial \chi_n^* / \partial n_e)_d \delta n_{e,d} + \sum_i (\partial \chi_n^* / \partial n_i)_d \delta n_i^* \quad (7-48)$$

with a similar expression for $\delta \eta_{dn}^*$; in equation (7-48), $\partial / \partial T$ applies to the explicit appearances of T in $\exp(-hv/kT)$, $\alpha_{xx}(v, T)$, etc., and similarly for $\partial / \partial n_e$. Now from equation (5-35) we have relations of the form

$$\delta n_i^* = (\partial n_i^* / \partial T)_N|_d \delta T_d + (\partial n_i^* / \partial N)_T|_d \delta N_d \quad (7-49)$$

and similarly for $\delta n_{e,d}$, so that the linearized values of ρ_d , χ_{dn}^* and η_{dn}^* can all be collapsed down to expressions of the form

$$\delta \chi_{dn}^* = \left(\frac{\partial \chi_n^*}{\partial T} \right)_{N|d} \delta T_d + \left(\frac{\partial \chi_n^*}{\partial N} \right)_{T|d} \delta N_d \quad (7-50)$$

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

$$\sum_{d'=d-1}^{d+1} T_{dd',n} \delta J_{d'n} + \sum_{d'=d-1}^{d+1} U_{dd',n} \delta N_{d'} + \sum_{d'=d-1}^{d+1} V_{dd',n} \delta T_{d'} = K_d \quad (7-51)$$

Similarly, the linearized constraint of radiative equilibrium [equation (7-13)] is

$$\begin{aligned} \sum_n w_n (\chi_{dn} - n_{e,d} \sigma_e) \delta J_{dn} + \sum_n w_n (J_{dn} \delta \chi_{dn} - \delta \eta_{dn} - \sigma_e J_{dn} \delta n_{e,d}) \\ = \sum_n w_n [\eta_{dn} - (\chi_{dn} - n_{e,d} \sigma_e) J_{dn}] \end{aligned} \quad (7-52)$$

and hydrostatic equilibrium [equation (7-9)] yields

$$\begin{aligned} (4\pi/c) \sum_n w_n (f_{dn} \delta J_{dn} - f_{d-1,n} \delta J_{d-1,n}) \\ + k(T_d \delta N_d + N_d \delta T_d - T_{d-1} \delta N_{d-1} - N_{d-1} \delta T_{d-1}) \\ = g(m_d - m_{d-1}) - N_d k T_d + N_{d-1} k T_{d-1} \\ - (4\pi/c) \sum_n w_n (f_{dn} J_{dn} - f_{d-1,n} J_{d-1,n}) \end{aligned} \quad (7-53)$$

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

$$\delta J_n \equiv (\delta J_{1n}, \delta J_{2n}, \dots, \delta J_{Dn})^T, \quad (n = 1, \dots, N) \quad (7-54)$$

$$\delta T \equiv (\delta T_1, \delta T_2, \dots, \delta T_D)^T \quad (7-55)$$

$$\delta N \equiv (\delta N_1, \delta N_2, \dots, \delta N_D)^T \quad (7-56)$$

Then equations (7-39), (7-52) and (7-53) yield

$$\begin{pmatrix} T_1 & 0 & \cdots & 0 & U_1 & V_1 \\ 0 & T_2 & & \ddots & U_2 & V_2 \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ 0 & & \cdots & 0 & \ddots & \vdots \\ W_1 & W_2 & \cdots & W_N & U_N & V_N \\ X_1 & X_2 & \cdots & X_N & C & D \end{pmatrix} \cdot \begin{pmatrix} \delta J_1 \\ \delta J_2 \\ \vdots \\ \delta J_N \\ \delta N \\ \delta T \end{pmatrix} = \begin{pmatrix} K_1 \\ K_2 \\ \vdots \\ K_N \\ L \\ M \end{pmatrix} \quad (7-57)$$

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 bi-diagonal. 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 δJ_n from the n th "row" into the last two "rows," one frequency at a time. Thus we solve for

$$\delta J_n + (T_n^{-1}U_n)\delta N + (T_n^{-1}V_n)\delta T = T_n^{-1}K_n \quad (7-58)$$

and eliminate δJ_n to obtain a final system of the form

$$\begin{pmatrix} P & Q \\ R & S \end{pmatrix} \begin{pmatrix} \delta N \\ \delta T \end{pmatrix} = \begin{pmatrix} F \\ G \end{pmatrix} \quad (7-59)$$

$$\text{where } P \equiv A - \sum_n W_n T_n^{-1} U_n; \quad Q \equiv B - \sum_n W_n T_n^{-1} V_n \\ R \equiv C - \sum_n X_n T_n^{-1} U_n; \quad S \equiv D - \sum_n X_n T_n^{-1} V_n \\ F \equiv L - \sum_n W_n T_n^{-1} K_n; \quad \text{and } G \equiv M - \sum_n X_n T_n^{-1} K_n$$

The final system (7-59) is solved for δN and δT .

Using δN and δT to revise the density and temperature, we can at each point m_d solve for new values of all the $n_{e,d}^*$'s and $n_{e,d}$ [using equations (5-27) through (5-31) iterated to consistency] and hence new values of χ_{dn}^* and η_{dn}^* , which are, in turn, used to obtain a formal solution of the transfer equation for new values of J_{dn} and f_{dn} ($d = 1, \dots, D$; $n = 1, \dots, N$). We use these new estimates to reconstruct equation (7-57), and iterate; as the solution improves, K_n , L , and M all $\rightarrow 0$, hence δN and $\delta T \rightarrow 0$.

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

$$(\partial x / \partial T)_{p_g} \equiv (\partial x / \partial T)_N + (\partial x / \partial N)(\partial N / \partial T)_{p_g},$$

and then explicitly assume $\delta p_g \equiv 0$. This eliminates the last "row" of system (7-57), and we solve only for δT ; 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 Δr 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 $(\Delta\rho)_E = (dp/dr)_A \Delta r$; 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 $(dp/dr)_R$ is the density gradient in the radiative surroundings, instability occurs if

$$(\Delta\rho)_E = (dp/dr)_A \Delta r < (\Delta\rho)_R = (dp/dr)_R \Delta r \quad (7-60)$$

(recall that $dp/dr < 0$). 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 $(dp/dr)_A$ and $(dp/dr)_R$, and demanding the pressure gradients be equal, we find from equation (7-60) that the Schwarzschild condition for instability is

$$[(\gamma - 1)/\gamma](-d \ln p/dr)_R < (-d \ln T/dr)_R \quad (7-61)$$

or $\nabla_R \equiv (d \ln T/d \ln p)_R > (\gamma - 1)/\gamma = (d \ln T/d \ln p)_A \equiv \nabla_A \quad (7-62)$

In stellar atmospheres the gas is not perfect because of the effects of ionization and radiation pressure; we may account for this by generalizing γ to Γ (160, 57) and writing $\nabla_A = (\Gamma - 1)/\Gamma$ where Γ 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 Γ , 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 ∇_A ,

and hence the critical value of ∇_R at which convection occurs. Thus for a perfect monatomic gas $\nabla_A = \frac{2}{3}/\frac{5}{3} = 0.4$, while for pure radiation pressure $\Gamma = \frac{4}{3}$, so $\nabla_A = 0.25$, and for conditions where hydrogen is ionizing Γ may be only 1.1 so ∇_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 $(-dT/dr) = (3\pi F \chi_R)/(16\sigma_R T^3)$, which implies (from hydrostatic equilibrium) $\nabla_R = (3\pi F \chi_R p)/(16\sigma_R g \rho T^4)$. 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 Γ 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: ∇_R is the radiative gradient that would occur if convection were suppressed; ∇_A is the adiabatic gradient; ∇_E is the gradient of the convective elements; and ∇ 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

$$\nabla_R \geq \nabla \geq \nabla_E \geq \nabla_A \quad (7-63)$$

Consider now a rising element of material. If δT 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 Δr with an average speed \bar{v} , the energy flux transported is

$$\pi F_{\text{conv}} = \rho C_p \bar{v} \delta T = \rho C_p \bar{v} [(-dT/dr) - (-dT/dr)_E] \Delta r \quad (7-64)$$

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

$$\pi F_{\text{conv}} = \frac{1}{2} \rho C_p \bar{v} T (\nabla - \nabla_E)(l/H) \quad (7-65)$$

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 μ 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

$$f_b = (gQ\rho/T) \delta T = (gQ\rho/T)[(-dT/dr) - (-dT/dr)_E] \Delta r \quad (7-66)$$

The buoyancy force is thus linear in the displacement; integrating over a total displacement Δ , 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

$$\bar{w} = \int_0^\Delta f_b(\Delta r) d(\Delta r) = (gQ\rho H/8)(\nabla - \nabla_E)(l/H)^2 \quad (7-67)$$

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

$$\bar{v} = (gQH/8)^{\frac{1}{2}}(\nabla - \nabla_E)^{\frac{1}{2}}(l/H) \quad (7-68)$$

and, therefore, from equation (7-65),

$$\pi F_{\text{conv}} = (gQH/32)^{\frac{1}{2}}(\rho C_p T)(\nabla - \nabla_E)^{\frac{1}{2}}(l/H)^2 \quad (7-69)$$

One of the uncertainties 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 ∇ and ∇_E in terms of ∇_R and ∇_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

$$\gamma = \frac{\text{excess energy content at time of dissolution}}{\text{energy lost by radiation during lifetime of element}} \quad (7-70)$$

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

$$\gamma = (\nabla - \nabla_E)/(\nabla_E - \nabla_A) \quad (7-71)$$

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 δT , 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

$$\begin{aligned} \gamma_{\text{thin}} &= (\rho C_p V \delta T) / [4\pi(4\sigma_R T^3/\pi)(\delta T/2)(\bar{\chi}_R V)(l/\bar{v})] \\ &= (\rho C_p \bar{v}) / (8\sigma_R T^3 \tau_e) \end{aligned} \quad (7-72)$$

where τ_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 δT , by writing $(-dT/dr) \approx (\delta T/l)$. Assuming a surface area A and the same lifetime as before, we now have

$$\gamma_{\text{thick}} = (\rho C_p V \delta T) / [(16\sigma_R T^3/3\bar{\chi}_R)(\delta T/l)A(l/\bar{v})] = (\rho C_p \bar{v} / 16\sigma_R T^3) 3\bar{\chi}_R (V/A) \quad (7-73)$$

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

$$\gamma_{\text{thick}} = \frac{1}{2}\tau_e(\rho C_p \bar{v})/(8\sigma_R T^3) \quad (7-74)$$

We interpolate between the two extreme cases by writing

$$\gamma = [(\rho C_p \bar{v})/(8\sigma_R T^3)] \cdot [(1 + \frac{1}{2}\tau_e^2)/\tau_e] \quad (7-75)$$

Combining equations (7-71) and (7-75), and substituting equation (7-68) for \bar{v} we derive finally

$$\frac{\nabla_E - \nabla_A}{(\nabla - \nabla_E)^{\frac{1}{2}}} = \frac{16\sqrt{2}\sigma_R T^3}{\rho C_p(gQH)^{\frac{1}{2}}(l/H)} \frac{\tau_e}{(1 + \frac{1}{2}\tau_e^2)} \equiv B \quad (7-76)$$

The final requirement we place upon the theory is that the correct total flux be transported by radiation and convection together—i.e.,

$$\pi F = \pi F_{\text{rad}} + \pi F_{\text{conv}} = \sigma_R T_{\text{eff}}^4 \quad (7-77)$$

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 ∇ , $\nabla_R \geq \nabla \geq \nabla_A$, 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

$$A(\nabla - \nabla_E)^{\frac{1}{2}} = \nabla_R - \nabla \quad (7-78)$$

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

$$A(\nabla - \nabla_E)^{\frac{1}{2}} + (\nabla - \nabla_E) + B(\nabla - \nabla_E)^{\frac{1}{2}} = (\nabla_R - \nabla_A) \quad (7-79)$$

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 $T(\tau_R)$ relation (adjusted to match the current values of T and τ_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 ∇ will be close to ∇_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 ∇ will lie close to ∇_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_{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_{rad} and F_{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 ($4000^{\circ}\text{K} \leq T_{\text{eff}} \leq 50,000^{\circ}\text{K}$, $2 \leq \log g \leq$

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 $\{\mu_i, v_i\}$ we can write

$$4\pi \sum_{i=1}^I w_i \mu_i^2 (u_{d+1,i} - u_{di}) / \Delta \tau_{d+\frac{1}{2},i} + \pi F_{\text{conv}}|_{d+\frac{1}{2}} = \sigma_R T_{\text{eff}}^4 \quad (7-80)$$

The convective flux can be regarded as $F_{\text{conv}}(p, p_g, T, \nabla)$ [given these variables, ∇_E follows from equation (7-76) and F_{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

$$F_{\text{conv}} = F_{\text{conv}}^0 + (\partial F_c / \partial p_g) \delta p_g + (\partial F_c / \partial T) \delta T + (\partial F_c / \partial \nabla) \delta \nabla \quad (7-81)$$

are computed numerically. Several approximations are introduced (274) to reduce this to an expression in δT only, and practical procedures for handling numerical problems have been developed (274; 275). Improvements in convergence might be obtained by including terms in δN as well as δT , 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

$$8000^\circ\text{K} \leq T_{\text{eff}} \leq 50,000^\circ\text{K}, \quad 2 \leq \log g \leq 5$$

(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_{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_{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_v/F_{v_0} , where v_0

denotes some prechosen frequency. In a few cases it is possible to make the comparison in *absolute* energy units using fluxes in $\text{ergs cm}^{-2} \text{ sec}^{-1} \text{ Hz}^{-1}$ 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* ($3650 \text{ \AA} \leq \lambda \leq 8205 \text{ \AA}$) 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*,

$$D_B \equiv 2.5 \log[F_v(\lambda 3650^+)/F_v(\lambda 3650^-)]$$

and the *Paschen jump*, $D_P \equiv 2.5 \log[F_v(\lambda 8205^+)/F_v(\lambda 8205^-)]$. 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 measurable, 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 standardized 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_{eff} and are insensitive to surface gravity (see Figure 7-3); hence both may be used to infer T_{eff} .

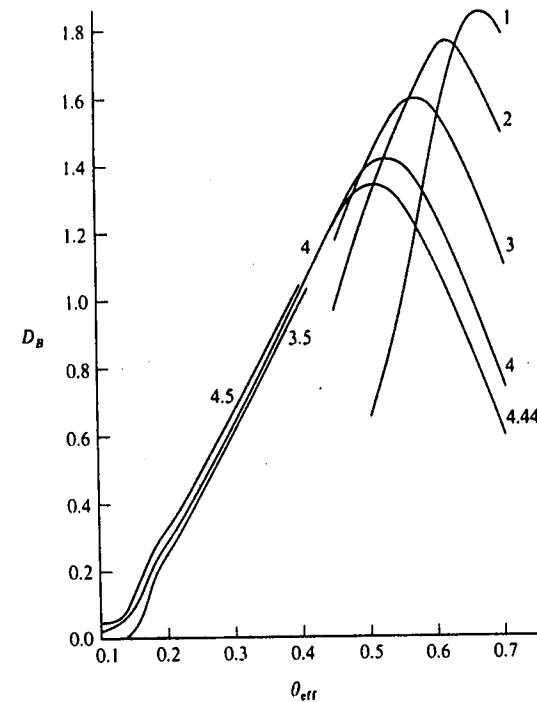


FIGURE 7-3
Balmer jumps computed from LTE model atmospheres, as a function of effective temperature and gravity. Ordinate: Balmer jump in magnitudes; abscissa: $\theta_{\text{eff}} \equiv 5040/T_{\text{eff}}$. Curves are labeled with $\log g$.

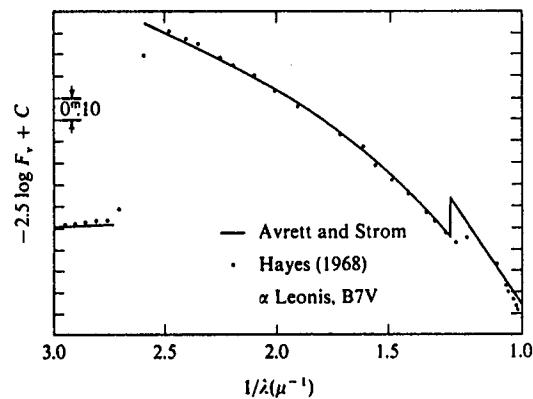


FIGURE 7-4
Comparison of the energy distribution of α Leo (B7V) as observed by Hayes (285) with the model atmosphere (608) that best fits the Paschen continuum, namely, with $T_{\text{eff}} = 13,000^{\circ}\text{K}$ and $\log g = 4$. Note that the computed and observed Balmer jumps are consistent. *Ordinate:* relative flux in magnitudes; *abscissa:* $1/\lambda$ where λ 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_{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

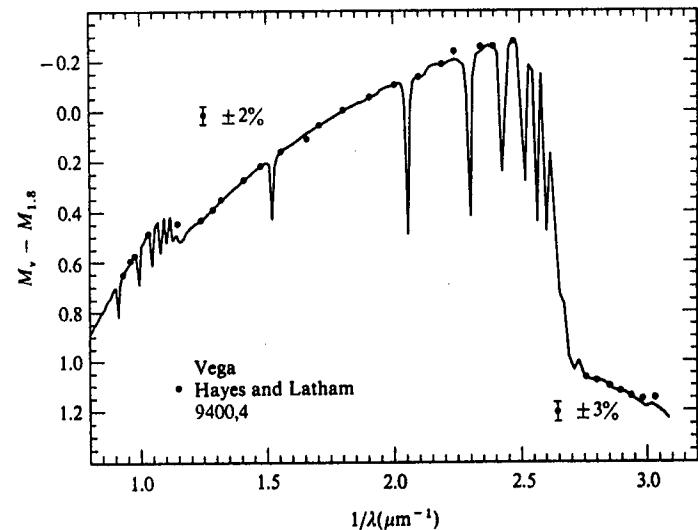


FIGURE 7-5
Comparison of energy distribution of the fundamental standard Vega as measured by Hayes and Latham (287) (dots) with a line-blanketed model [(381; 561, 271)] that has $T_{\text{eff}} = 9400^{\circ}\text{K}$ and $\log g = 4$. *Ordinate:* relative flux in magnitudes; *abscissa:* $1/\lambda$ where λ is in microns.

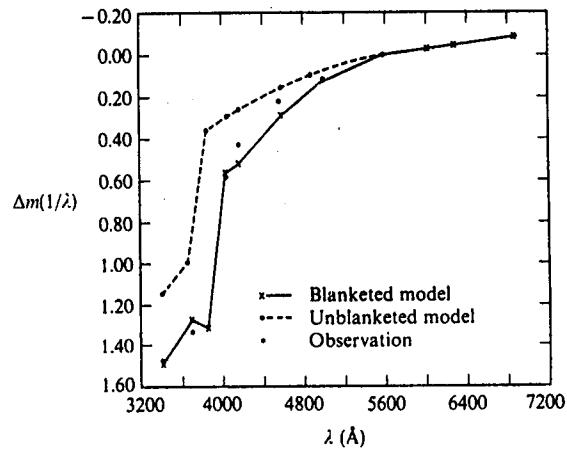


FIGURE 7-6
Unblanketed and blanketed energy distributions for models with $T_{\text{eff}} = 6500^{\circ}\text{K}$ and $\log g = 4$, compared with observations of Procyon. *Ordinate:* relative flux in magnitude units; *abscissa:* wavelength λ in Å. From (612), by permission.

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 $f_v = 3.50 \times 10^{-20} \text{ ergs cm}^{-2} \text{ sec}^{-1} \text{ hz}^{-1}$ at $\lambda 5556 \text{ \AA}$. For any other star we use the magnitude difference Δm 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_{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_{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 $912 \text{ \AA} \leq \lambda \leq 1600 \text{ \AA}$ by the "direct" approach. The effects of the blanketing are quite dramatic. The integrated flux of the blanketed model corresponds to $T_{\text{eff}} = 21,900^\circ\text{K}$, 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 $T_{\text{eff}} = 24,000^\circ\text{K}$. 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,

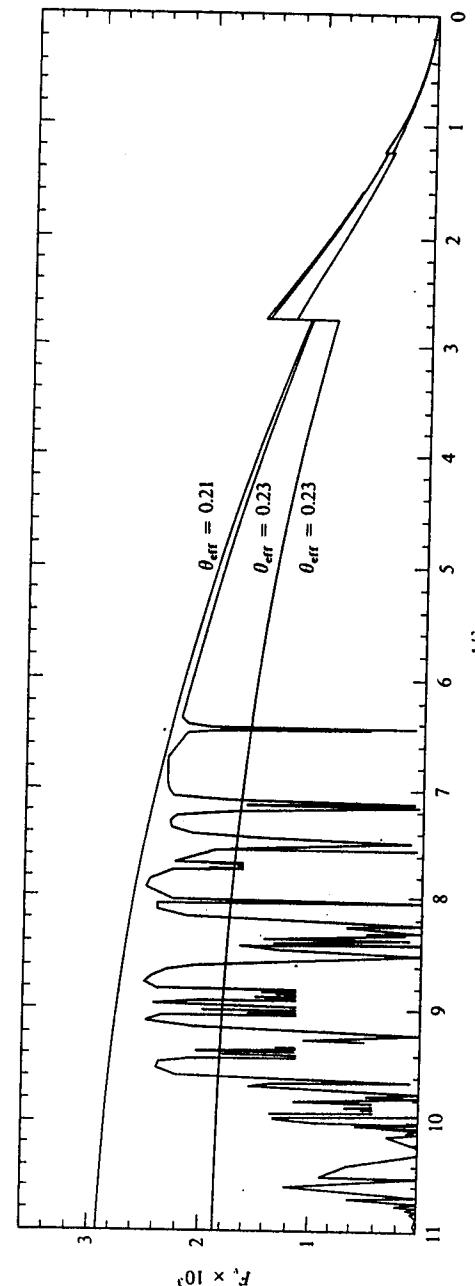


FIGURE 7-7
Flux from blanketed and unblanketed LTE models (449). The blanketed model yields a total energy output corresponding to $T_{\text{eff}} = 21,900^\circ\text{K}$ ($\theta_{\text{eff}} = 0.23$), but because the ultraviolet flux has been redistributed into the visible, the flux there is much higher than that of an unblanketed model of the same temperature, and, in fact, matches an unblanketed model with $T_{\text{eff}} = 24,000^\circ\text{K}$ ($\theta_{\text{eff}} = 0.21$).
Abscissa: $1/\lambda$, where λ is in microns; ordinate: $F_v \times 10^3 \text{ ergs cm}^{-2} \text{ sec}^{-1} \text{ hz}^{-1}$. From (449), by permission.

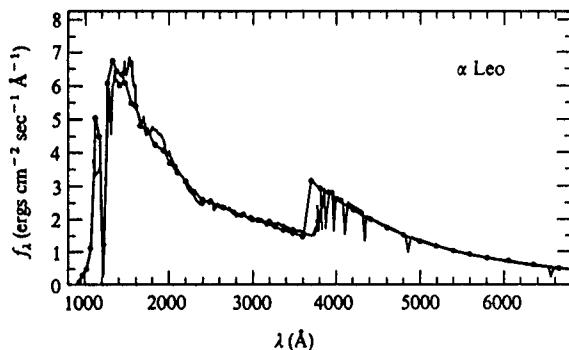


FIGURE 7-8
Comparison of empirical absolute energy distribution of α Leo (B7V) (516, 221) with a line-blanketed model (381) of the same effective temperature (12,200°K). The agreement is excellent and lends strong support to the model techniques. *Ordinate:* absolute flux $10^9 f_\lambda$ in ergs $\text{cm}^{-2} \text{sec}^{-1} \text{\AA}^{-1}$ at the earth; *abscissa:* wavelength λ in \AA .

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_{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

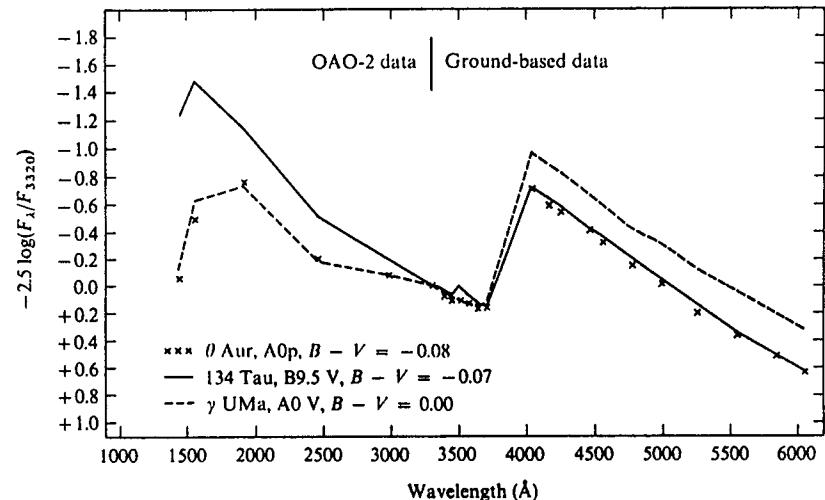


FIGURE 7-9
Comparison of relative energy distribution of the peculiar (Si 3995) A-star θ Aur with those of 134 Tau (B9.5V) and γ U Ma (A0V) (391). Because of the enhanced line blanketing arising from the greater heavy-element abundances in the peculiar star, the energy distribution of θ Aur matches neither 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 \AA . From (391), by permission.

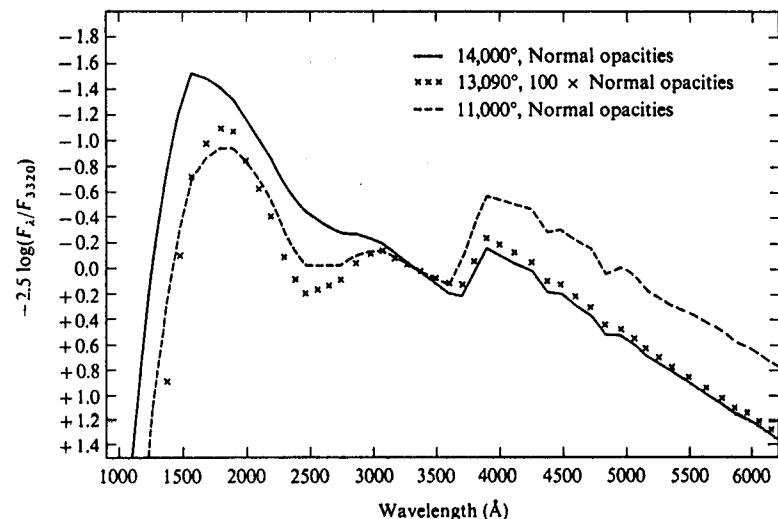


FIGURE 7-10
Line-blanketed models (391) showing effects of 100-fold enhanced heavy-element abundances; note strong resemblance to the effects shown in Figure 7-9. From (391), by permission.

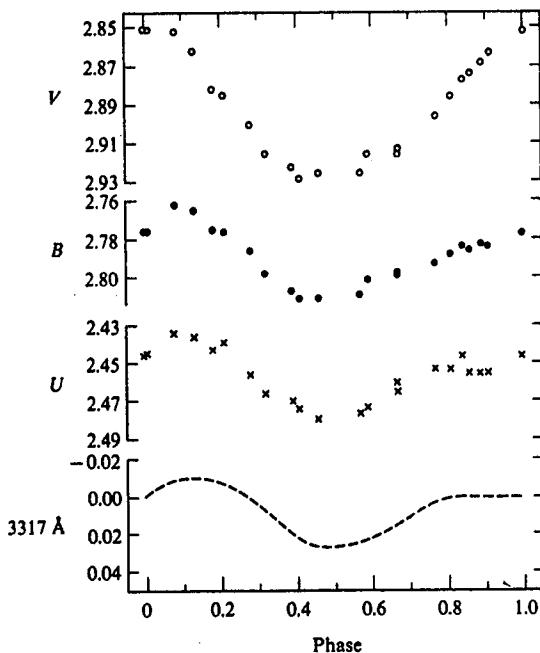


FIGURE 7-11
Variation of the peculiar (Si-Cr-Eu) A-star α^2 CVn in UBV and at $\lambda 3317 \text{ \AA}$ as measured from OAO-2. From (464), by permission.

the peculiar star has a *lower* T_{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_{eff} . A further effect is shown in the Ap (Si-Cr-Eu) spectrum variable α^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

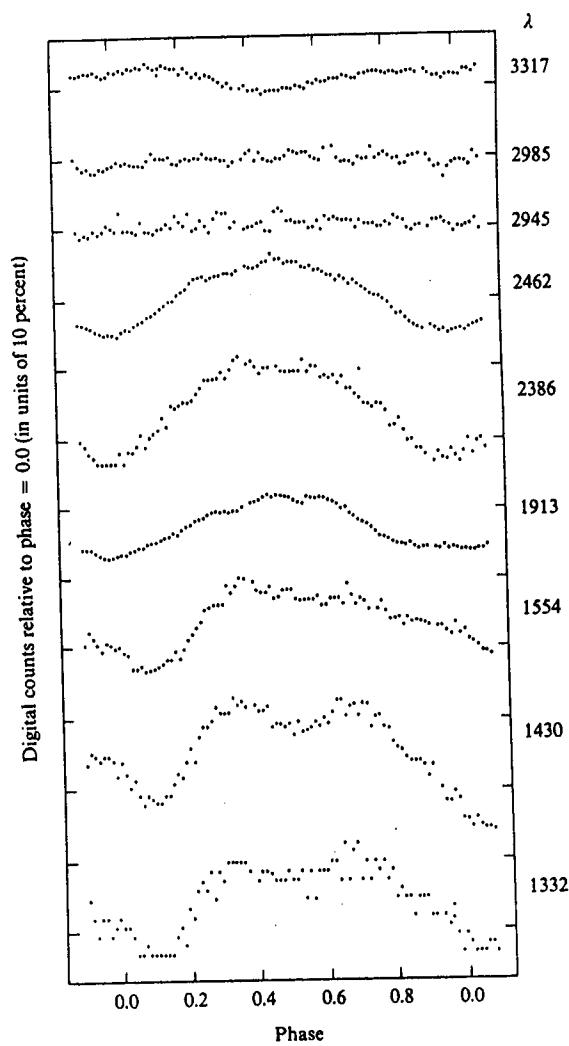


FIGURE 7-12
Variation of α^2 CVn in ultraviolet as measured from OAO-2. Curves are labeled with wavelengths (\AA) of filters. Note *antiphase* variation of far ultraviolet flux relative to visible! From (464), by permission.

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

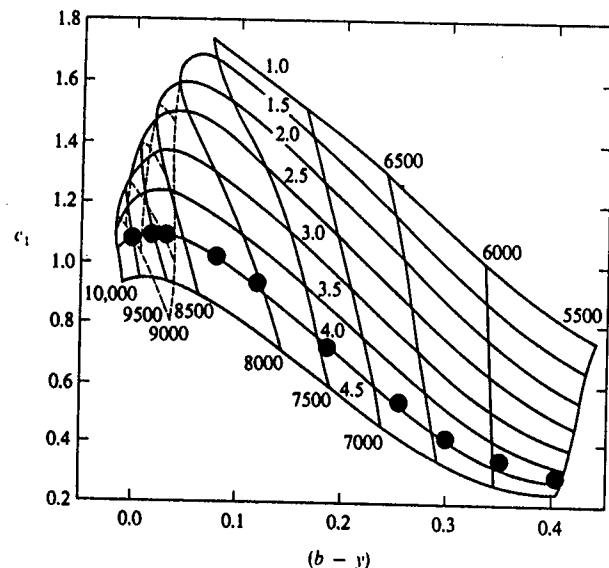


FIGURE 7-13
Comparison of observed (dots) Strömgren (c_1 , $b - y$) indices for main-sequence stars (which have $\log g \approx 4$) [(516, 17; 516, 45)] with calculated values from line-blanketed models [(381; 516, 271)]; note good agreement in $\log g$, which suggests that T_{eff} can be estimated accurately.

we have

$$(i - j)_{\text{obs}} = -2.5 \log \left[\int_0^{\infty} T_i(\lambda) F(\lambda) d\lambda / \int_0^{\infty} T_j(\lambda) F(\lambda) d\lambda \right] + k_{ij} \quad (7-82)$$

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_{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 indispensable for estimating T_{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_{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, $\chi_v \equiv \kappa_c + \sigma$, and the emissivity is given by $\eta_v = \kappa_c B_v + \sigma J_v$; then the condition of radiative equilibrium is

$$\kappa_c \int_0^{\infty} B_v(T_0) dv = \kappa_c \int_0^{\infty} J_v^0 dv \quad (7-83)$$

Here the scattering term has cancelled out. From the Eddington-Barbier relation, we expect that at the surface (i.e., $\tau_v \ll 1$), $J_v^0 \approx B_v(\tau_v \approx 1)$, hence $J_v^0 \approx \frac{1}{2}B_v(T_{\text{eff}})$. 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 $v \leq v_0$, $\kappa = \gamma\kappa_c$ for $v > v_0$. Then the atmosphere will equilibrate to some new surface temperature T'_0 given by

$$\kappa_c \int_0^{v_0} B_v(T'_0) dv + \gamma\kappa_c \int_{v_0}^{\infty} B_v(T'_0) dv = \kappa_c \int_0^{v_0} J_v dv + \gamma\kappa_c \int_{v_0}^{\infty} J_v dv \quad (7-84)$$

Assuming that for $v \leq v_0$, $J_v \approx J_v^0$ (i.e., neglecting backwarming), and noting that for $v > v_0$, the surface value of $J_v \approx \frac{1}{2}B_v(T'_0)$, equation (7-84) can be rewritten as

$$\begin{aligned} \kappa_c \int_0^{\infty} B_v(T'_0) dv &\approx \kappa_c \int_0^{\infty} J_v^0 dv \\ &\quad - \kappa_c \left[(\gamma - 1) \int_{v_0}^{\infty} B_v(T'_0) dv + \int_{v_0}^{\infty} J_v^0 dv - \gamma \int_{v_0}^{\infty} J_v dv \right] \\ &\approx \kappa_c \int_0^{\infty} J_v^0 dv \\ &\quad - \kappa_c \left\{ \frac{\gamma}{2} \int_{v_0}^{\infty} B_v(T'_0) dv + \int_{v_0}^{\infty} \left[\frac{1}{2} B_v(T_{\text{eff}}) - B_v(T'_0) \right] dv \right\} \end{aligned} \quad (7-85)$$

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 γ . This result is not rigorous because we expect J_v must, in fact, rise above J_v^0 for $v \leq v_0$; but we shall see below that a rigorous analysis verifies the correctness of our conclusion. Suppose we evaluate $T'(\tau_v)$ at some point inside the atmosphere where $\tau_v \gg 1$ for $v > v_0$ while $\tau_v < 1$ for $v < v_0$. 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 $T'(\tau_v)$ 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, $\chi_v = \kappa_c + \sigma + \sum_i l_i \phi_{v_i}$, and make both a thermal and scattering contribution to the emissivity, $\eta_v = \kappa_c B_v + \sigma J_v + \sum_i l_i \phi_{v_i} [\epsilon_i B_v + (1 - \epsilon_i) J_v]$. Then the condition of radiative equilibrium reduces to

$$\begin{aligned} \kappa_c \int_0^{\infty} B_v(T'_0) dv &= \kappa_c \int_0^{\infty} J_v^0 dv \\ &\quad - \left\{ \sum_i l_i \epsilon_i [B_{v_i}(T'_0) - J_i] + \kappa_c \sum_i \int_{\Delta_i} (J_v^0 - J_v) dv \right\} \end{aligned} \quad (7-86)$$

where J_i denotes $\int \phi_{v_i} J_v dv$ for the i th line, and Δ_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 ϵ_i . In LTE, $\epsilon_i \equiv 1$, $J \approx \frac{1}{2}B_v(T'_0)$, 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., $\epsilon_i \rightarrow 0$), 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 Chandrasekhar (150) and further developed by Münch (474). In this model we assume (a) the continuum opacity is frequency-independent, (i.e., $\kappa_v \equiv \kappa$); (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

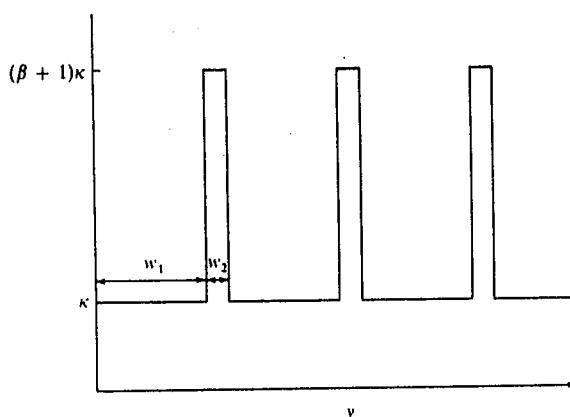


FIGURE 7-14
Picket-fence model. Lines are assumed to be a factor of β more opaque than continuum, and to occur with a probability $w_2 = 1 - w_1$ in any frequency band.

frequencies in the continuum,

$$\mu(dI_v^{(1)}/d\tau) = I_v^{(1)} - B_v \quad (7-87a)$$

and, in the line

$$\mu(dI_v^{(2)}/d\tau) = (1 + \beta)I_v^{(2)} - (1 - \varepsilon)\beta J_v^{(2)} - (1 + \varepsilon\beta)B_v \quad (7-87b)$$

Integrating over all frequencies (radiation quantities without a subscript v denote integrated quantities) and accounting for the relative probabilities that the band is covered by line or continuum, we find

$$\mu(dI^{(1)}/d\tau) = I^{(1)} - w_1 B \quad (7-88a)$$

$$\text{and} \quad \mu(dI^{(2)}/d\tau) = (1 + \beta)I^{(2)} - (1 - \varepsilon)\beta J^{(2)} - (1 + \varepsilon\beta)w_2 B \quad (7-88b)$$

These equations are to be solved simultaneously with a constraint of radiative equilibrium, which is obtained by integrating over angle and demanding that $F^{(1)} + F^{(2)} = \text{constant}$, namely

$$J^{(1)} + (1 + \varepsilon\beta)J^{(2)} = [w_1 + w_2(1 + \varepsilon\beta)]B \quad (7-89)$$

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

$$\mu(dI^{(l)}/d\tau) = \gamma_l(I^{(l)} - w_l B), \quad (l = 1, 2) \quad (7-90)$$

where, from equation (7-89),

$$B = \sum_{l=1}^2 \gamma_l J^{(l)} / \sum_{l=1}^2 w_l \gamma_l \quad (7-91)$$

To solve this system we use the discrete-ordinate approach, and choose $\{\mu_i\}$, ($i = \pm 1, \dots, \pm n$), such that

$$J^{(l)} = \frac{1}{2} \sum_{j=-n}^n a_j I_j^{(l)} \quad (7-92)$$

Then, substituting equations (7-91) and (7-92) into (7-90), we have

$$\frac{\mu_l}{\gamma_l} \frac{dI_l^{(l)}}{d\tau} = I_l^{(l)} - \left(\frac{w_l}{2 \sum_m w_m \gamma_m} \right) \sum_{m=1}^2 \gamma_m \sum_{j=-n}^n a_j I_j^{(m)}, \quad (l = 1, 2) \\ (i = \pm 1, \dots, \pm n) \quad (7-93)$$

If we now assume a solution of the form

$$I_l^{(l)} = C w_l e^{-k_l \tau} / (1 + k_l \mu_i / \gamma_l) \quad (7-94)$$

we find that k satisfies the characteristic equation

$$\sum_{m=1}^2 w_m \gamma_m = \sum_{m=1}^2 w_m \gamma_m \sum_{j=1}^n a_j / (1 - k^2 \mu_j^2 / \gamma_m^2) \quad (7-95)$$

This equation yields $2n - 1$ nonzero roots for k^2 (bounded by poles at $1/\mu_1^2, \dots, 1/\mu_n^2$ and $\gamma^2/\mu_1^2, \dots, \gamma^2/\mu_n^2$) 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

$$I_l^{(l)} = b w_l (\tau + Q + \mu_l / \gamma_l) \quad (7-96)$$

which may be verified by direct substitution into equation (7-93). The general solution (7-93) is thus of the form

$$I_l^{(l)}(\tau) = w_l b \left(\tau + Q + \frac{\mu_l}{\gamma_l} + \sum_{\alpha=1}^{2n-1} \frac{L_\alpha e^{-k_\alpha \tau}}{1 + k_\alpha \mu_l / \gamma_l} + \sum_{\alpha=1}^{2n-1} \frac{L_{-\alpha} e^{k_\alpha \tau}}{1 - k_\alpha \mu_l / \gamma_l} \right), \\ (l = 1, 2) \\ (i = \pm 1, \dots, \pm n) \quad (7-97)$$

Demanding that the solution not diverge exponentially as $\tau \rightarrow \infty$, we set $L_{-\alpha} \equiv 0$ for all α . Requiring that the total flux

$$F = 2 \sum_{l=1}^2 \sum_{j=-n}^n a_j \mu_j I_j^{(l)} \quad (7-98)$$

$$\text{we find} \quad b = \left(\frac{3}{4} F \right) / \sum_{l=1}^2 w_l \gamma_l^{-1} \quad (7-99)$$

The constant Q and the L_α 's are determined from the surface boundary conditions $I_{-l}^{(l)}(0) \equiv 0$, which yield a linear system of $2n$ equations in $2n$ unknowns:

$$Q - (\mu_l / \gamma_l) + \sum_{\alpha=1}^{2n-1} L_\alpha / (1 - k_\alpha \mu_l / \gamma_l) = 0, \quad (l = 1, 2) \\ (i = 1, \dots, n) \quad (7-100)$$

Using equations (7-99), (7-97), and (7-92), we find

$$J^{(l)}(\tau) = \frac{\frac{3}{4} F w_l}{\sum w_m \gamma_m^{-1}} \left(\tau + Q + \sum_{\alpha=1}^{2n-1} L_\alpha e^{-k_\alpha \tau} \sum_{j=1}^n \frac{a_j}{1 - k_\alpha^2 \mu_j^2 / \gamma_l^2} \right) \quad (7-101)$$

and, from equation (7-91),

$$B(\tau) = \frac{3}{4} F \left(\tau + Q + \sum_{\alpha=1}^{2n-1} L_\alpha e^{-k_\alpha \tau} \right) / \sum_{m=1}^2 w_m \gamma_m^{-1} \quad (7-102)$$

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 τ_R exceeds τ ; in the limit of infinitely strong lines ($\gamma_2 \rightarrow \infty$), $\tau_R(\tau) = \tau/w_1$, 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

$$S(x) = Q - x + \sum_{a=1}^{2n-1} L_a / (1 - k_a x) \quad (7-103)$$

The boundary conditions (7-100) show that $S(x) = 0$ at the $2n$ values $x = \mu_i/\gamma_i$. 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., $R(x) \equiv \prod_{a=1}^{2n-1} (1 - k_a x)$], 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 $R(x)S(x) = C(x - \mu_1) \cdots (x - \mu_n)(x - \mu_1/\gamma_1) \cdots (x - \mu_n/\gamma_n)$. If we equate the coefficients of the two terms in x^n on the two sides of this equation we can evaluate C as $C = k_1 k_2 \cdots k_{2n-1}$, and hence obtain finally

$$S(x) = k_1 \cdots k_{2n-1} \left[\prod_{l=1}^2 \prod_{i=1}^n (x - \mu_i/\gamma_i) \right] / \prod_{a=1}^{2n-1} (1 - k_a x) \quad (7-104)$$

which implies that

$$S(0) = k_1 k_2 \cdots k_{2n-1} \mu_1^2 \mu_2^2 \cdots \mu_n^2 / \gamma^n \quad (7-105)$$

Now consider the characteristic function

$$\begin{aligned} T(X) &= \sum_{m=1}^2 w_m \gamma_m \left[1 - \sum_{j=1}^n a_j / (1 - \mu_j^2 / \gamma_m^2 X) \right] \\ &= \sum_{m=1}^2 w_m \gamma_m \left[1 - X \sum_{j=1}^n a_j / (X - \mu_j^2 / \gamma_m^2) \right] \\ &= \sum_{m=1}^2 w_m \gamma_m^{-1} \sum_{j=1}^n a_j \mu_j^2 [(\mu_j^2 / \gamma_m^2) - X]^{-1} \end{aligned} \quad (7-106)$$

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

$$T(X) = \frac{1}{3} \left(\sum_{m=1}^2 w_m \gamma_m^{-1} \right) \prod_{k=1}^{2n-1} (X_k - X) / \left\{ \prod_{m=1}^2 \prod_{j=1}^n [(\mu_j^2 / \gamma_m^2) - X] \right\} \quad (7-107)$$

From the middle equality of equation (7-106), we have $T(0) = \sum w_m \gamma_m$; and from equation (7-107), we have

$$T(0) = \frac{1}{3} (\sum w_m \gamma_m^{-1}) / [(\mu_1^2 \cdots \mu_n^2 k_1 \cdots k_{2n-1}) / \gamma^n]^2$$

Combining these two results we then find, from equation (7-105),

$$S(0) = \left(\sum_{m=1}^2 w_m \gamma_m^{-1} \right)^{\frac{1}{2}} \left(3 \sum_{m=1}^2 w_m \gamma_m \right)^{-\frac{1}{2}} \quad (7-108)$$

But comparison of equations (7-102) and (7-103) shows that

$$B(0) = \frac{3}{4} F S(0) / (\sum w_m \gamma_m^{-1}) \quad (7-109)$$

Hence we conclude that

$$[B(0)/F] = (\sqrt{3}/4) [\sum w_m \gamma_m] (\sum w_m \gamma_m^{-1})^{-\frac{1}{2}} \quad (7-110)$$

This result may be restated in a form that reveals its physical content. The Planck mean opacity is

$$\bar{\kappa}_P \equiv B^{-1} \int_0^\infty \kappa_v B_v dv = B^{-1} \kappa (w_1 B + w_2 \gamma B) = \kappa (w_1 + w_2 \gamma) \quad (7-111)$$

while the Rosseland mean opacity is

$$\begin{aligned} (\bar{\kappa}_R)^{-1} &= (dB/dT)^{-1} \int_0^\infty \kappa_v^{-1} (dB_v/dT) dv \\ &= (dB/dT)^{-1} \kappa^{-1} (w_1 + w_2/\gamma) (dB/dT) \end{aligned} \quad (7-112)$$

$$\text{or} \quad \bar{\kappa}_R = \kappa (w_1 + w_2/\gamma)^{-1} \quad (7-113)$$

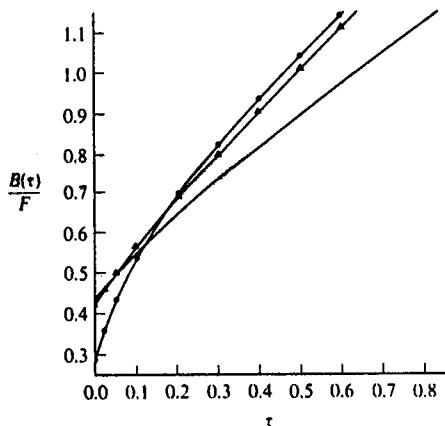


FIGURE 7-15
Depth-variation of integrated Planck function in picket-fence models. Solid curve: grey solution, $\beta = 1$; solid dots: $y = 10, w_1 = 0.8, w_2 = 0.2, \epsilon = 1$ (LTE); triangles: $y = 10, w_1 = 0.8, w_2 = 0.2, \epsilon = 0$ (pure scattering). Note backwarming effect in both blanketed 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

$$[B(0)/F] = (\sqrt{3}/4)(\bar{\kappa}_R/\bar{\kappa}_P)^{\frac{1}{2}} \quad (7-114)$$

or

$$(T_0/T_{\text{eff}}) = (\sqrt{3}/4)^{\frac{1}{2}}(\bar{\kappa}_R/\bar{\kappa}_P)^{\frac{1}{2}} \quad (7-115)$$

Now in the limit as $\gamma \rightarrow \infty$, the Rosseland mean (being a reciprocal mean) simply saturates at a value κ/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 $B(t)/F$ is plotted for the grey case and for one of Münch's solutions with $\epsilon = 1, w_1 = 0.8, w_2 = 0.2$, and $y = 10$; in this case $B(0)/F$ decreases from the grey value 0.4330 to 0.286; i.e., T_0/T_{eff} drops from 0.811 to 0.721.

The analysis just described can also be applied to an opacity jump at a critical frequency, v_0 , beyond which the opacity increases by a factor of γ . We now apply the two versions of equation (7-90) on the ranges ($v \leq v_0$) and ($v \geq v_0$), and define $w_1 B \equiv \int_{v_0}^{\infty} B_v dv$ and $w_2 B \equiv \int_{v_0}^{\infty} B_v dv$. We must then assume that w_1 and w_2 are constant with depth; for example we may

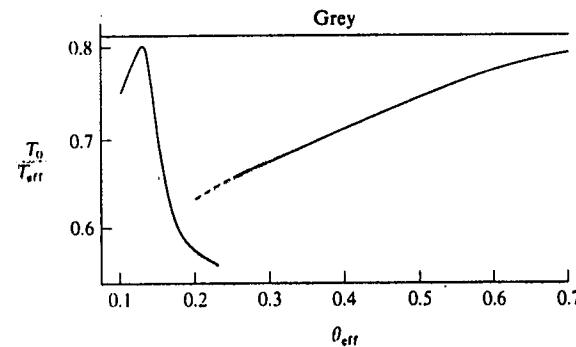


FIGURE 7-16
Ratio of boundary temperature T_0 to effective temperature T_{eff} as a function of $\theta_{\text{eff}} \equiv 5040/T_{\text{eff}}$. The break near $\theta_{\text{eff}} = 0.25$ results from inclusion of the Lyman continuum in the high-temperature models. Upper line gives value of T_0/T_{eff} 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_{eff} from nongrey LTE model-atmospheres calculations. For all $\theta_{\text{eff}} \geq 0.25$, the Lyman continuum has been omitted. For the coolest models T_0/T_{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_{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_{eff} rises toward the grey value again. At still higher temperatures T_0/T_{eff} drops again as a result of the He I edge at $\lambda 504 \text{ \AA}$ and the He II edge at $\lambda 226 \text{ \AA}$.

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 $n^* \alpha_{v_i} (b - f)$ over all bound levels with $u_n \equiv n^{-2}(\chi_{\text{ion}}/kT) \leq u \equiv (hv/kT)$, 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

$$\kappa_v^* = Cu^{-3}(1 - e^{-u}) \left[1 + \sum_{u > u} 2u_1 n^{-3} \exp(u_1/n^2) \right] \quad (7-116)$$

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 $B_v = C'u^3e^{-u}(1 - e^{-u})^{-1}$,

$$\bar{\kappa}_P(u_0) = C'' \left\{ 1 - e^{-u_0} + \sum_n 2u_1 n^{-3} [1 - \exp(-u_0 + n^{-2}u_1)] \right\} \quad (7-117)$$

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

$$\begin{aligned} \bar{\kappa}_P(\infty)/\bar{\kappa}_P(u_1) &= \left(1 + 2u_1 \sum_{n=1}^{\infty} n^{-3} \right) / \left(1 + 2u_1 \sum_{n=2}^{\infty} n^{-3} \right) \\ &= (1 + 2.4u_1)/(1 + 0.4u_1) \end{aligned} \quad (7-118)$$

For $u_1 = 7.2$ we thus find $\bar{\kappa}_P(\infty)/\bar{\kappa}_P(u_1) = 4.7$, hence

$$T_0(\text{Lyman cont.})/T_0(\text{no Lyman cont.}) = (4.7)^{-1} = 0.825 \quad (7-119)$$

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 $T_{\text{eff}} = 15,000^{\circ}\text{K}$, $\log g = 4$, 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 $T \approx 10,300^{\circ}\text{K}$ for $-4 \leq \log \tau \leq -2$ 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

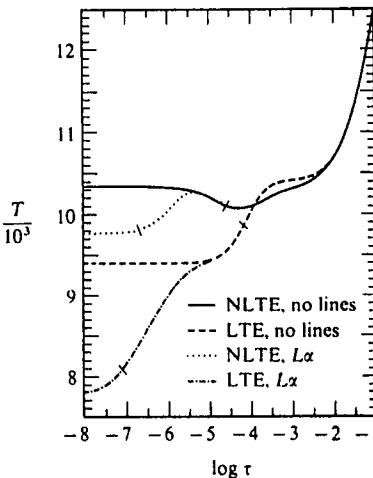


FIGURE 7-17
Temperature distribution for LTE and non-LTE models with $T_{\text{eff}} = 15,000^{\circ}\text{K}$ and $\log g = 4$. The atmosphere is composed of hydrogen, which is represented by a schematic model atom with two bound levels and continuum. This model atom accounts for the Lyman, Balmer, and free-free continua, and the Lyman- α 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

$$\mu(dI^{(1)}/d\tau) = I^{(1)} - w_1\sigma(J^{(1)} + \lambda J^{(2)}) \quad (7-88a')$$

$$\text{and} \quad \mu(dI^{(2)}/d\tau) = \gamma I^{(2)} - (\gamma - w_1\sigma\lambda)J^{(2)} - w_2\sigma\lambda J^{(1)} \quad (7-88b')$$

Applying the discrete-ordinate method, we obtain the characteristic equation (474)

$$1 - w_1\sigma G - (1 - w_1\sigma\lambda\gamma^{-1})H + w_1\sigma(1 - \lambda\gamma^{-1})GH = 0 \quad (7-120)$$

$$\text{where} \quad G \equiv \frac{1}{2} \sum_{i=-n}^n a_i/(1 + k\mu_i) \quad (7-121a)$$

$$\text{and} \quad H \equiv \frac{1}{2} \sum_{i=-n}^n a_i/(1 + k\mu_i/\gamma) \quad (7-121b)$$

Equation (7-120) has $2n - 1$ positive roots k_α . The solution for $B(\tau)$ is

$$B(\tau) = \frac{3}{4} F \left(\tau + Q + \sum_{\alpha=1}^{2n-1} M_\alpha e^{-k_\alpha \tau} \right) / \sum_m w_m \gamma_m^{-1} \quad (7-122)$$

where $M_a \equiv \sigma L_a [G_a(1 - G_a)^{-1} + (\lambda/\gamma)H_a(H_a - 1)^{-1}]$ and where, in turn, the $2n - 1$ constants L_a and the constant Q are determined from the boundary conditions $I_{-i}^{(m)}(0) \equiv 0$, which imply

$$Q + w_1^{-1} \sum_{a=1}^{2n-1} L_a(1 - G_a)^{-1}(1 - k_a\mu_i)^{-1} = \mu_i, \quad (i = 1, \dots, n)$$

(7-123a)

and

$$Q + (yw_2)^{-1} \sum_{a=1}^{2n-1} L_a(H_a - 1)^{-1}(1 - k_a\mu_i/\gamma)^{-1} = (\mu_i/\gamma), \quad (i = 1, \dots, n)$$

(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 $\chi_v = \chi_v^* + \delta\chi_v$. If we define $b_i \equiv n_i/n_i^*$, then $\delta\chi_v = \sum_i d_i b_i n_i^* \alpha_{ik}(v)$ 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 δT and δJ_k set to zero. The latter method yields a consistent current value for n_e and n_{ion} ; the former does not, for it ignores the non-linearity 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

$$\mu(dI_v/dz) = -(\chi_v^* + \delta\chi_v) I_v + \kappa_v^* B_v + n_e \sigma_e J_v \quad (7-124)$$

$$\text{or} \quad \mu(dI_v/d\tau_v) = I_v - \xi_v B_v - \zeta_v J_v \quad (7-125)$$

where $\xi_v \equiv \kappa_v^*/(\chi_v^* + \delta\chi_v)$ and $\zeta_v \equiv n_e \sigma_e/(\chi_v^* + \delta\chi_v)$, 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

$$\begin{aligned} d_i \left[4\pi \int_{v_0}^{\infty} (\alpha_v J_v / hv) dv + \sum_{j \neq i}^{\kappa} C_{ij} \right] - \sum_{j \neq i}^{\kappa} d_j C_{ij} \\ = 4\pi \int_{v_0}^{\infty} (B_v - J_v)(1 - e^{-hv/kT})(\alpha_v/hv) dv \quad (7-126) \end{aligned}$$

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_{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 ζ_v and ζ_v in equation (7-125) (specifically, $\zeta_v \neq 1 - \zeta_v$). 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 $T_{\text{eff}} = 10,000^{\circ}\text{K}$, $\log g = 3$, 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 $n \geq 3$ 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 $B_v \sim \exp(-h\nu/kT)$ rises rapidly. Thus at the surface $J_v \approx \frac{1}{2}B_v(T_{\text{eff}})$ exceeds $B_v(T_0)$, and the level is preferentially photoionized as can be seen from equation (7-126).

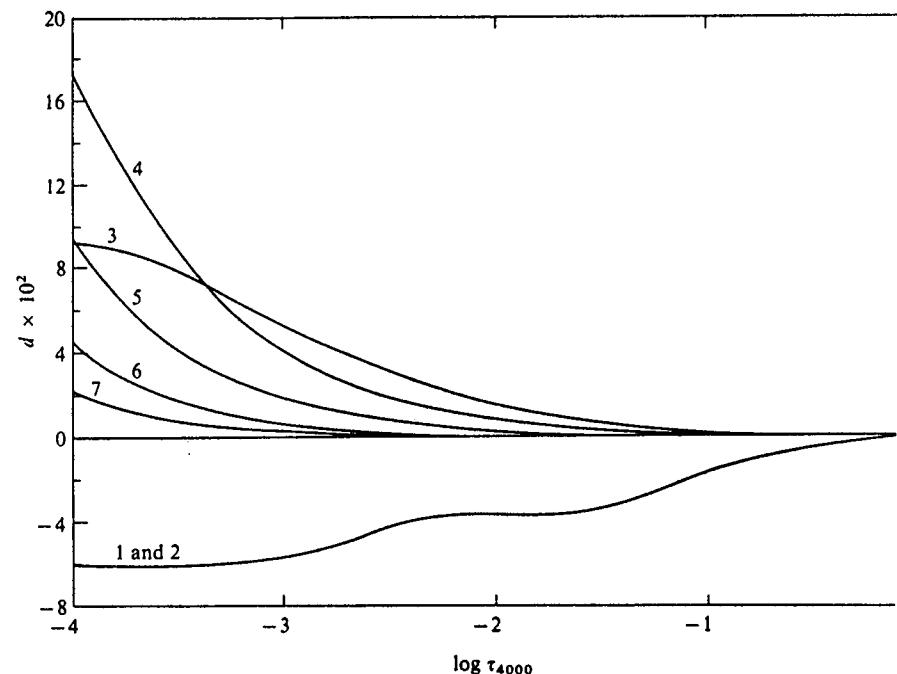


FIGURE 7-18
Non-LTE departure coefficients in the first 7 states of hydrogen for a model with $T_{\text{eff}} = 10,000^{\circ}\text{K}$ and $\log g = 3$. Ordinate gives $d_i \times 10^2$; curves are labeled with quantum number of the level of the model atom. Note that level 2 is *underpopulated* while 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 $n \geq 3$, $h\nu/kT \lesssim 1$, and the dilution factor of $\frac{1}{2}$ in J_v outweighs the effects of the temperature gradient, so $J_v < B_v$, 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 $d_1 \sum_{j=2}^{\infty} C_{1j} = \sum_{j=2}^{\infty} d_j C_{1j}$ 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., $T_{\text{eff}} \gtrsim 35,000^{\circ}\text{K}$), 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 $dH_v/d\tau_v = J_v - S_v$). 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 $n_i R_{ik}$ to depart from $n_k R'_{ki}$, 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_{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 v_0 ; consider only frequencies $v \geq v_0$, and suppose that $hv_0/kT \gg 1$ 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" $\phi_v \equiv (v_0/v)^3$. Writing $n_i = b_i n_i^*$, we then have $\chi_v = \chi_0 \phi_v = (b_1 \chi_1^* + \chi_u^*) \phi_v$, where the superscript * denotes LTE values, and the subscript u denotes the contribution of the upper-level and free-free continua. Similarly, $\eta_v = (\eta_1^* + \eta_u^*) \phi_v = (\chi_1^* + \chi_u^*) B_v \phi_v$. Let $d\tau_0 = -\chi_0 dz$ be the optical depth at the continuum head. Then the transfer equation to be solved is

$$\mu(dI_v/d\tau_0) = \phi_v(I_v - S_v) \quad (7-127)$$

$$\text{where } S_v = (\chi_1^* + \chi_u^*) B_v / (b_1 \chi_1^* + \chi_u^*) = \{[(1-r)/b_1] + r\} B_v \quad (7-128)$$

$$\text{and } r \equiv \chi_u^* / (b_1 \chi_1^* + \chi_u^*) \quad (7-129)$$

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 τ_0), but in equation (7-128) we substitute an *analytical* expression for $(1/b_1)$ obtained from the ground-state statistical equilibrium equation:

$$(1/b_1) = \left[4\pi \int_{v_0}^{\infty} (\alpha_v/hv) J_v dv + C \right] / \left[4\pi \int_{v_0}^{\infty} (\alpha_v/hv) B_v dv + C \right] \quad (7-130)$$

where $C \equiv C_{12} + C_{1k}$. Substituting equation (7-130) into (7-128) shows that S_v is of the form

$$S_v = \gamma_v \int \Phi_v J_v dv + \epsilon_v B_v, \quad (7-131)$$

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 $hv/kT \gg 1$, the frequency variation of J_v and B_v [roughly as $\exp(-hv/kT)$] shows so rapid a drop with increasing v that practically all of the contribution to the photoionization and recombination rate integrals comes from $v \approx v_0$. Therefore, we replace the integrals by $4\pi w_0(\alpha_0/hv_0)J_0$ and $4\pi w_0(\alpha_0/hv_0)B_0$,

where w_0 is an appropriate weight factor, and consider the transfer problem only at the continuum head, $v = v_0$ [see also (195)]. If we define $\varepsilon \equiv C/[4\pi w_0(\alpha_0/hv_0)B_0 + C]$ then

$$S_0 = (1 - r)[(1 - \varepsilon)J_0 + \varepsilon B_0] + rB_0 \approx (1 - \bar{\varepsilon})J_0 + \bar{\varepsilon}B_0 \quad (7-132)$$

where $\bar{\varepsilon} \equiv \varepsilon + r$, and the transfer equation in the Eddington approximation is

$$\frac{1}{3}(d^2J_0/d\tau_0^2) = \bar{\varepsilon}(J_0 - B_0) \quad (7-133)$$

Assuming $\bar{\varepsilon}$ is constant, and that the atmosphere is nearly isothermal so that B_0 is roughly constant, the solution of equation (7-133) can be written immediately [cf. equation (6-9)] as

$$J_0 = B_0\{\bar{\varepsilon}^{\frac{1}{3}} + 1 - \exp[-\sqrt{3\bar{\varepsilon}}\tau_0]\}/(1 + \bar{\varepsilon}^{\frac{1}{3}}) \quad (7-134)$$

which shows that (a) the solution *thermalizes* only at depths $\gtrsim 1/\bar{\varepsilon}^{\frac{1}{3}}$, and (b) the departure coefficient at the surface is $b_1(0) \approx \bar{\varepsilon}^{-\frac{1}{3}}$. Because $\bar{\varepsilon} \ll 1$, the *departure from LTE is large, and persists to great depths*. The actual value of $\bar{\varepsilon}$ is a bit difficult to estimate if collisions dominate (i.e., $\varepsilon > r$), for the density, and hence ε , increases exponentially into the atmosphere; on the other hand, the parameter r is almost depth-independent, and often dominates. For example, at $T \sim 25,000^\circ\text{K}$, $r \sim 10^{-2}$ [see (633, 193)] and $b_1 \sim 10$, which is in agreement with calculation. In summary we see that it is important to use the information inherent in the statistical equilibrium equations directly in the transfer equation; further, given the dominant scattering term, it is obvious that one must use care in satisfying the constraint of radiative equilibrium.

We now shall consider a *partial linearization method* that is fairly general, and provides a satisfactory means of solving the non-LTE *continuum-formation* problem, with energy balance; the method fails, however, when lines are included, in which case one must use the complete linearization method described in the next subsection. The basic idea in the method [cf. (40; 41)] is to incorporate the statistical equilibrium equations into the transfer equations by manipulating them analytically to obtain explicit expressions for source functions, and to incorporate the constraint of radiative equilibrium into the transfer problem by linearization. Then, in principle, one solves all three sets of equations: transfer, statistical and radiative equilibrium, simultaneously.

The transfer equation is

$$d^2(f_v J_v)/d\tau_v^2 = J_v - (\eta_v + n_e \sigma_e J_v)/\chi_v = J_v - S_v \quad (7-135)$$

where (omitting bound-bound transitions)

$$\chi_v = \sum_i (n_i - n_i^* e^{-hv/kT}) \alpha_{ik}(v) + \sum_\kappa n_e n_\kappa \alpha_{\kappa k}(v, T) (1 - e^{-hv/kT}) + n_e \sigma_e \quad (7-136)$$

$$\begin{aligned} \text{and } \eta_v &= (2hv^3/c^2) e^{-hv/kT} \left[\sum_i n_i^* \alpha_{ik}(v) + \sum_\kappa n_e n_\kappa \alpha_{\kappa k}(v, T) \right] \\ &= B_v \left[\sum_i n_i^* \alpha_{ik}(v) + \sum_\kappa n_e n_\kappa \alpha_{\kappa k}(v, T) \right] (1 - e^{-hv/kT}) = \kappa_v^* B_v \end{aligned} \quad (7-137)$$

We now divide the spectrum into a series of characteristic frequency ranges ($v_{i0} \leq v \leq v_{i1}$) upon which we assume the opacity and emissivity of a particular continuum $i \rightarrow \kappa$ is much larger than all others. For example we might assume that the Lyman continuum dominates for all frequencies such that $\lambda \leq 912 \text{ \AA}$, the Balmer continuum dominates for $912 \text{ \AA} \leq \lambda \leq 3650 \text{ \AA}$, the free-free continuum dominates for $\lambda \geq \lambda_l$, the threshold of the last bound level, etc. We then factor out the dominant terms from both the emissivity and the opacity, and write, for ($v_{i0} \leq v \leq v_{i1}$),

$$(\eta_v/\chi_v) = [n_i^*/(n_i - n_i^* e^{-hv/kT})] \xi_{iv} B_v = \xi_{iv} B_v / (b_i - e^{-hv/kT}) \quad (7-138)$$

where the three quantities

$$\xi_{iv} \equiv [(n_i - n_i^* e^{-hv/kT})/n_i^*](\kappa_v^*/\chi_v) \quad (7-139)$$

f_v , and τ_v are all evaluated using *current values* of the temperature and level populations, whereas the term $(b_i - e^{-hv/kT})$ in equation (7-138) is to be replaced with an *analytical expression*, involving the radiation field, to be found from the statistical equilibrium equations. The statistical equilibrium equation for level i is [cf. equation (7-126)]

$$b_i \left(R_{ik} + \sum_{j \neq i}^\kappa C_{ij} \right) = R_{ki} + C_{ik} + \sum_{j \neq i}^I b_j C_{ij} \quad (7-140)$$

where R_{ik} and R_{ki} are defined, as usual, by equations (5-66) and (5-67). Solving for $b_i - e^{-hv/kT}$ we then have

$$\begin{aligned} (b_i - e^{hv/kT})^{-1} &= \left(R_{ik} + \sum_{j \neq i}^\kappa C_{ij} \right) / \left[R_{ki} - e^{-hv/kT} R_{ik} \right. \\ &\quad \left. + \sum_{j \neq i}^I (b_j - e^{-hv/kT}) C_{ij} + C_{ik} (1 - e^{-hv/kT}) \right] \end{aligned} \quad (7-141)$$

so that, defining $\zeta_v \equiv n_e \sigma_e / \chi_v$ (again using current values), the source function in equation (7-135) on the range ($v_{i0} \leq v \leq v_{i1}$) can be written as

$$S_v = \xi_{iv} \left(\gamma_{iv} \int_{v_{i0}}^{v_{i1}} \Phi_{iv} J_v dv + \varepsilon_{iv} B_v \right) + \zeta_v J_v \quad (7-142)$$

where

$$\Phi_{iv} \equiv 4\pi \alpha_{ik}(v)/hv$$

$$\gamma_{iv} \equiv B_v/D_{iv}$$

$$\varepsilon_{iv} \equiv \left(\int_{v_{i1}}^{\infty} \Phi_{iv} J_v dv + \sum_{j \neq i}^k C_{ij} \right) / D_{iv} \quad (7-143)$$

and D_{iv} is the term in square brackets in the denominator of equation (7-141). In the evaluation of γ_{iv} and ε_{iv} we again use current values for the b_j 's of other levels and for the terms in J_v and T ; it is thus clear that a large amount of information is *lagged* in this technique. Note that ε_{iv} contains an *overlap integral* from those frequency ranges outside the presumed range of dominance (v_{i0}, v_{i1}) of the transition $i \rightarrow k$; usually this integral is small compared to the term displayed explicitly in the scattering integral. Equation (7-142) accounts for the most important terms directly, but because the information involving other levels and in the nonlinear terms is lagged, iteration will still be required. Finally, discretizing, we have equations of the form

$$d^2(f_k J_k) / d\tau_k^2 = J_k - \xi_{ik} \left(\gamma_{ik} \sum_{l=k_{i0}}^{k_{i1}} w_l \Phi_{il} J_l + \varepsilon_{ik} B_k \right) - \zeta_k J_k \quad (7-144)$$

on the range ($k_{i0} \leq k \leq k_{i1}$) delineated for the i th transition.

We must now solve the system (7-144) subject to the constraint of radiative equilibrium which requires that

$$\begin{aligned} \int_0^{\infty} \chi_v (J_v - S_v) dv &= \sum_{k=1}^K w_k \chi_k J_k - \sum_{l=1}^I \sum_{k=k_{i0}}^{k_{i1}} w_k \chi_k \\ &\times \left[\xi_{ik} \left(\gamma_{ik} \sum_{l=k_{i0}}^{k_{i1}} w_l \Phi_{il} J_l + \varepsilon_{ik} B_k \right) + \zeta_k J_k \right] = 0 \quad (7-145) \end{aligned}$$

We suppose that B_v , the value of the Planck function that *does* yields radiative equilibrium, can be written in terms of the current value B_v^0 as $B_v = B_v^0 + (\partial B_v^0 / \partial T) \Delta T$. Substituting into equation (7-145) we find

$$\Delta T = \sum_{k=1}^K w_k \psi_k J_k - E_1 \quad (7-146)$$

where

$$\psi_k \equiv \left[\chi_k (1 - \zeta_k) - \Phi_{ik} \sum_{l=k_{i0}}^{k_{i1}} w_l \chi_l \xi_{il} \gamma_{il} \right] / E_2, \quad (k_{i0} \leq k \leq k_{i1}) \quad (7-147)$$

$$E_1 \equiv \left(\sum_{l=1}^I \sum_{k=k_{i0}}^{k_{i1}} w_k \chi_k \xi_{ik} \varepsilon_{ik} B_k^0 \right) / E_2 \quad (7-148)$$

$$\text{and } E_2 \equiv \sum_{l=1}^I \sum_{k=k_{i0}}^{k_{i1}} w_k \chi_k \xi_{ik} \varepsilon_{ik} (\partial B^0 / \partial T)_k \quad (7-149)$$

Thus equation (7-144) becomes

$$\begin{aligned} d^2(f_k J_k) / d\tau_k^2 &= (1 - \zeta_k) J_k \\ &- \xi_{ik} \left[\gamma_{ik} \sum_{l=k_{i0}}^{k_{i1}} w_l \Phi_{il} J_l + \varepsilon_{ik} (\partial B^0 / \partial T)_k \sum_{k'=1}^K w_{k'} \psi_{k'} J_{k'} \right] \\ &- \xi_{ik} \varepsilon_{ik} [B_k^0 - (\partial B^0 / \partial T)_k E_1] \quad (7-150) \end{aligned}$$

which is completed with boundary conditions given by equations (7-37a) and (7-37c) (the latter specifies the flux explicitly).

Equation (7-150) has the following important properties. (a) It incorporates both the equations of statistical equilibrium and radiative equilibrium in terms of the *new* (i.e., as yet undetermined) radiation field; thus a solution of the system satisfies both of these constraints automatically. (b) Cancellation of large scattering terms is done *analytically* in the *coefficients* [cf. especially equation (7-147)] of the radiation field before the transfer equation is solved; hence good control of the residual thermal terms is obtained. (c) The equations are of the standard Feautrier form [equation (6-31)] with the frequency coupling spread over the entire spectrum, and can be solved using the usual elimination scheme of equations (6-39) through (6-41). (Why is the Feautrier method more efficient than the Rybicki scheme in this problem, even though the number of frequencies is large?)

The whole solution proceeds iteratively. After solving equation (7-150) for J_v , the temperature distribution is revised via equation (7-146), the new radiative rates are used to solve the statistical equilibrium equations for new occupation numbers, new Eddington factors are then calculated from the revised values of η_v and χ_v , and the procedure starting with equation (7-139) is repeated.

Methods of the type described above (which is related to the *equivalent-two-level-atom* approach for line-formation problems; see §12-2) have been developed by Feautrier (211) and Auer and Mihalas (40; 41). Feautrier's

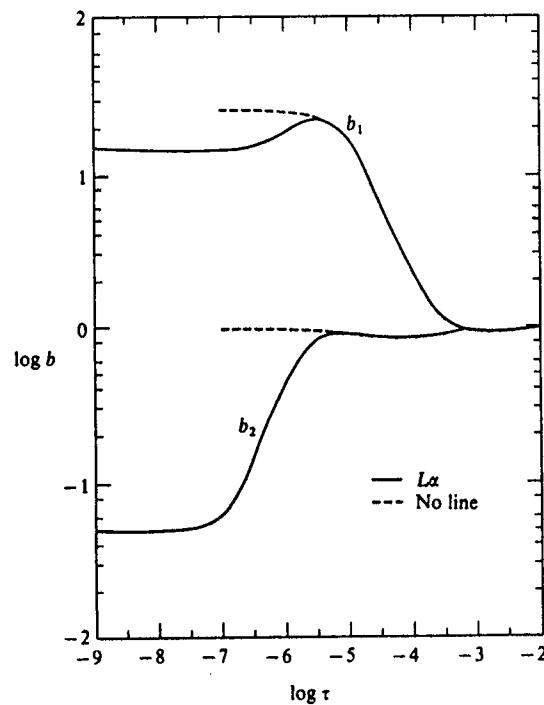


FIGURE 7-19
Non-LTE departure coefficients in the first two states of hydrogen for a model with $T_{\text{eff}} = 15,000^{\circ}\text{K}$ and $\log g = 4$. Dashed curves: Lyman, Balmer and free-free continua only; full curves: $L\alpha$ line included. The optical depth scale in the abscissa is measured just longward of the Balmer jump.
From (40), by permission.

formalism differs significantly in appearance from that displayed above, but the physical content is about the same. These methods have been used successfully to construct continuum models for B-stars, including the Lyman continuum in cases where lambda iteration would be utterly hopeless. Characteristically the ground state of hydrogen becomes strongly overpopulated in the outer layers; see Figure 7-19. This has important consequences for the temperature structure and emergent spectrum of the atmosphere, as will be discussed below.

Attempts have been made (40; 41) to apply the techniques described above to cases including line transitions. It is essential to include the lines because (a) they drastically affect both occupation numbers and energy balance in

the atmosphere, and (b) we obtain reliable level-populations that can be used to compute line profiles (to be compared with observation) *only* if line transitions have been taken into account in the solution of the statistical equilibrium equations. The importance of these effects is illustrated in Figures 7-19 and 7-20. In Figure 7-19 we see that the effect of including $L\alpha$ is to *drain* the $n = 2$ level, which goes from being slightly overpopulated to markedly underpopulated. The drop in b_1 is caused by the drop in temperature; actual occupation numbers n_1 do not change much [see (40)]. In Figure 7-20 we see a similar effect for b_2 and b_3 when $H\alpha$ is taken into account. If we had computed the $H\alpha$ profile with the continuum-only values (which are similar to those of Figure 7-18) we would have found a spurious emission core in the line (because $n = 3$ would have been too overpopulated relative

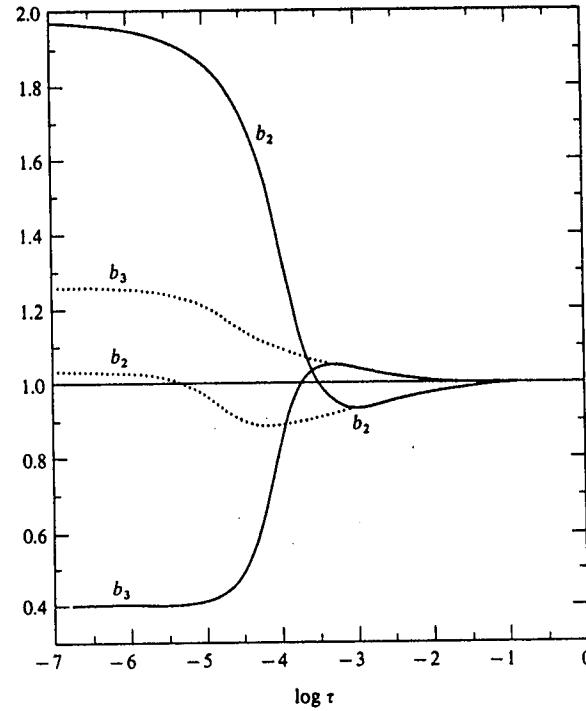


FIGURE 7-20
Non-LTE departure coefficients in the second and third level of hydrogen in a model with $T_{\text{eff}} = 15,000^{\circ}\text{K}$ and $\log g = 4$. Dotted curves: Lyman, Balmer, Paschen, and free-free continua only; full curves: $H\alpha$ included. Optical depth scale as in Figure 7-19.
From (41), by permission.

to $n = 2$); inclusion of the line from the outset drains $n = 3$ into $n = 2$ and produces a strong absorption line (we shall compare with observations in §12-5).

The above method does *not* work well for determining the temperature structure when lines are included; thus in (40), the solution for the temperature structure including $L\alpha$ tended to stabilize rather than converge, and in (41) a solution including $H\alpha$ (which is formed at about the same depths as the Lyman continuum) tended to become strongly unstable. The reasons for the failure can be traced to basic inadequacies of the method. (1) Many of the terms in the equations are lagged, and *interactions* among transitions are treated only iteratively rather than collectively from the outset. (2) The temperature dependence of many variables (e.g., the opacity) is ignored in the linearization procedure. (3) Most important, it is *presumed* that the temperature is, in some sense, "the fundamental variable" of the problem. Such a presumption is adequate for treating continua that are coupled strongly to the temperature via radiative recombination but, as noted before, is totally unsatisfactory for lines, where both emission and absorption rates are decoupled from the temperature structure. Let us now consider a method of great generality and power that overcomes these difficulties entirely.

THE COMPLETE LINEARIZATION METHOD

From a physical point of view, the solution of the non-LTE stellar atmospheres problem entails the specification, at each point in the medium, of the distribution of the radiation field as a function of frequency, the temperature and density of the material, and the distribution of atoms and ions over all bound states. These distribution functions are to be determined in such a way that the constraints of energy balance, momentum balance (i.e., hydrostatic equilibrium in the present work), steady-state statistical equilibrium, and charge and number conservation are all satisfied rigorously. To achieve the desired result, it is important to recognize two basic physical points, and to construct a generalized method in the light of them. First, *no one variable is more "fundamental" than any other, for they all interact*. Thus we must regard the solution at a given depth-point m_d to consist of the vector

$$\Psi_d = (J_1, \dots, J_K, N, T, n_e, n_1, \dots, n_L)^T, \quad (d = 1, \dots, D) \quad (7-151)$$

where D is the number of depth-points, K is the number of frequencies, and L is the total number of bound levels (of all kinds) considered. (Note that there is, in principle, redundant information in Ψ_d as written; e.g., given all the J 's, the n 's follow from the rate equations. However, all of the information is of interest in practice, and hence will be retained.) The coupling among all

these variables is intricate and essentially complete. For example, we have already seen in §5-5 that a change in the radiation field at *any* frequency implies a change in the occupation numbers of *all* bound states [cf. equation (5-108)]; further, any change in a radiation field implies a change in the local temperature and density, and so on. Second, *the variables interact globally throughout the atmosphere, and a change in any variable at a given point implies changes in all other variables at all other points*. Thus, if we alter occupation numbers locally, this changes the emissivity and opacity of the material, and hence, via the transfer equation, the radiation field throughout the atmosphere; when scattering terms dominate the source functions, information propagates over large distances in the material, and the global nature of the interaction is accentuated markedly. A truly adequate method must, therefore, put all variables on an equal footing, allow fully for all possible couplings among the physical variables resulting from the imposed physical constraints, and for the interaction of each variable with all others at *every* point in the atmosphere via transfer equations and requirements of energy and momentum balance.

To determine the solutions Ψ_d we require a total of $K + L + 3$ equations; we may choose these to be K transfer equations (J_k , $k = 1, \dots, K$); the equation of hydrostatic equilibrium (N); the equation of radiative equilibrium (T); a total particle-number conservation condition (n_e); and L statistical equilibrium equations specifying level-populations, total abundances of chemical species, and total charge conservation (n_1, \dots, n_L), exactly as written in §5-4 [see equations (5-91) through (5-93)]. The essential difficulty to be faced now is that these equations are *nonlinear* and must, in general, be solved by some kind of iteration procedure. In particular, we suppose that our desired (but as yet unknown) solution Ψ_d can be written in terms of the current (but imperfect) solution Ψ_d^0 as $\Psi_d = \Psi_d^0 + \delta\Psi_d$. We then choose $\delta\Psi_d$ so as to satisfy all constraints more closely; i.e., if $f_d(\Psi_d) = 0$ represents the entire system of constraints, we demand that $f_d(\Psi_d^0 + \delta\Psi_d) = 0$, and solve for $\delta\Psi_d$ by linearizing the entire system:

$$f_d(\Psi_d^0) + \sum_j \frac{\partial f_d}{\partial \Psi_{d,j}} \delta\Psi_{d,j} = 0 \quad (7-152)$$

If we express the transfer equations as difference equations, and the constraints in terms of quadrature sums, the f 's are coupled *algebraic* equations and the linearized system reduces to the standard block tridiagonal form of equation (6-31). In the step-by-step elimination scheme we account (consistently to first order) for the interaction of variables at one depth-point with all other depth-points, subject, ultimately, to the requirements of the imposed boundary conditions. The global nature of the problem is thus taken fully into account.

The complete linearization procedure described here is a generalization of that described in §7-2, and treats an enlarged set of constraints as well as a larger set of basic physical variables (42). As before, we introduce a depth discretization $\{m_d\}$ and frequency discretization $\{v_n\}$. We suppose that a starting solution ψ_d^0 , ($d = 1, \dots, D$), has already been obtained (assuming, say, LTE). The transfer equations are again equations (7-37a) through (7-37c) and the linearizations of these equations have already been written in equations (7-39) through (7-47) and in the results of Exercise 7-9. [See also (437, 22–32)]. The essential difference is that we must now use the general formulae of equations (7-1) and (7-2) for χ_v and η_v , and write

$$\delta\chi_{dn} = (\partial\chi_n/\partial T)_d \delta T_d + (\partial\chi_n/\partial n_e)_d \delta n_{e,d} + \sum_{l=1}^L (\partial\chi_n/\partial n_l)_d \delta n_{l,d} \quad (7-153)$$

and a similar expression for $\delta\eta_{dn}$ [detailed expressions for all these derivatives may be found in (437, 51–57)]. Further, we no longer regard the level populations as functions of (N, T) , but rather as independent variables whose coupling to T, J_v , etc. is specified by the rate equations.

For the hydrostatic and radiative equilibrium equations we again use equations (7-9) and (7-13) as linearized in equations (7-53) and (7-52), respectively, employing equation (7-153) for $\delta\chi_{dn}$ and a similar equation for $\delta\eta_{dn}$. The equation of particle number conservation can be written as

$$N_d = n_{e,d} + \sum_{l=1}^L n_{l,d} \quad (7-154)$$

If we assume the same hydrogen-helium mixture as described in §5-4 and let Y denote the (number) abundance of helium relative to hydrogen, we can rewrite equation (7-154) in a simpler form:

$$N_d = n_{e,d} + (1 + Y) \left[n_{p,d} + \sum_{l=1}^{L_H} n_{l,d}(H) \right] \quad (7-155)$$

which, when linearized, yields

$$\begin{aligned} -\delta N_d + \delta n_{e,d} + (1 + Y) \left[\delta n_{p,d} + \sum_{l=1}^{L_H} \delta n_{l,d}(H) \right] \\ = N_d - n_{e,d} - (1 + Y) \left[n_{p,d} + \sum_{l=1}^{L_H} n_{l,d}(H) \right] \end{aligned} \quad (7-156)$$

Finally, the rate equations and charge-conservation equations are of the form given in §5-4, and their linearized form [see equations (5-102) through

(5-104)] is

$$\delta n_d - (\partial n/\partial n_e)_d \delta n_{e,d} - (\partial n/\partial T)_d \delta T_d - \sum_{k=1}^K (\partial n/\partial J_k)_d \delta J_{dk} = 0 \quad (7-157)$$

where $\left(\frac{\partial n}{\partial x} \right) = \mathcal{A}^{-1} \left[\frac{\partial \mathcal{B}}{\partial x} - \left(\frac{\partial \mathcal{A}}{\partial x} \cdot \mathbf{n} \right) \right] \quad (7-158)$

for any variable x [see equations (5-105) through (5-108) and expressions in (42) and (437, 38–47)]. Here we assume that \mathbf{n} is already a solution of the current system $\mathcal{A}\mathbf{n} = \mathcal{B}$.

The complete system is of the general form

$$-\mathbf{A}_d \delta\psi_{d-1} + \mathbf{B}_d \delta\psi_d - \mathbf{C}_d \delta\psi_{d+1} = \mathbf{L}_d \quad (7-159)$$

which may be solved by the standard Feautrier elimination scheme.

Exercise 7-14: (a) Sketch the form of the \mathbf{A} , \mathbf{B} , and \mathbf{C} matrices, indicating nonzero elements by x 's as was done in §5-4 for the rate equations. Show that the \mathbf{A} and \mathbf{C} matrices are void below the row specifying hydrostatic equilibrium, and differ for that row [see (450, 130)]. (b) For typical problems, the number of depth-points $D \sim 70$, the number of frequencies $K \sim 100$, and the number of constraints $L + 3 \sim 15$. Show that despite the size of K , it is more economical to use the Feautrier elimination scheme than the Rybicki scheme.

In the system written above, \mathbf{L}_d is the residual error in the constraints found when current values ψ_d^0 are used; as $\mathbf{L}_d \rightarrow 0$, the corrections $\delta\psi_d \rightarrow 0$. We have already stressed the physical implications of the complete linearization. Mathematically, the equations are internally self-consistent, and are equivalent to a generalized Newton-Raphson procedure; convergence, if obtained, should therefore be quadratic. This is not achieved in practice because, after each set of corrections $\delta\psi$ have been determined and applied to the solution, it is necessary to recompute the Eddington factors (which were assumed fixed under linearization) from a formal solution using current source functions, as was done in the LTE case as well. Nevertheless, an order-of-magnitude reduction in the $\delta\psi$'s is often obtained from successive iterations, and on the whole the procedure is stable and efficient, and it easily handles physical problems that defeat the other techniques described above. It appears that at present the complete linearization method is the best technique available for solving non-LTE stellar-atmosphere problems. The method can handle multilevel, multiline, multispecies problems of great generality, and treats all the physical constraints because the full interaction

among all variables is allowed, and a priori assumptions about the dependence of quantities upon only a restricted set of variables (e.g., T) are avoided.

The complete linearization method has been used with good success to construct models in the temperature range from A-stars (42; 43; 368; 224) through O-stars (436; 45). The early work, carried out before the introduction of the variable Eddington factor technique (44) made use of the Eddington approximation (i.e., one angle-quadrature point) and has been supplanted by the later calculations. An extensive grid of models for O- and B-stars is available (430; 432). In this work, departures from LTE are taken into account for the first five levels of hydrogen, the first two levels of He I and He II, and for an "average light ion" (which represents C, N, and O) consisting of five stages of ionization, each with a ground-state only. These models typically allow for six hydrogen-line transitions: $L\alpha$, $L\beta$, $L\gamma$, $H\alpha$, $H\beta$, and $P\alpha$ for the O-stars, and $H\alpha$, $H\beta$, $H\gamma$, $P\alpha$, $P\beta$, and $B\alpha$ for the B-stars (for which the Lyman lines may be set in detailed balance), and yield results for the physical structure of the atmosphere, continuum parameters, and H-line profiles.

NON-LTE EFFECTS ON ENERGY DISTRIBUTIONS

Departures from LTE affect both the continuum and the line-spectrum from a stellar atmosphere; the discussion in this section focuses on the continuum only. Enough results are available for models of early-type stars to allow us to delineate the regions of the H-R diagram where departures from LTE have important effects and where they can be ignored. At the present time the non-LTE models do not include line-blanketing effects; thus we compare LTE and non-LTE *unblanketed* models to determine differential effects (which then, presumably, can be applied as corrections to values of parameters obtained from blanketed LTE models). Extensive sets of results for continuum jumps and Strömgren-system colors are given in (430; 516, 241), and complete energy distributions are given in (432); we shall summarize some of the principal results here, and we suggest the reader examine the literature cited for further detail.

For the B-stars ($15,000^{\circ}\text{K} \leq T_{\text{eff}} \leq 30,000^{\circ}\text{K}$) the effects of departures from LTE in the visible spectral regions are generally found to be negligible for main-sequence stars, but become important for giants and supergiants (low gravities). Non-LTE effects on colors are shown in Figure 7-21. There we see that results for both LTE and non-LTE models at a given gravity lie along the same curves, but the position of models with given T_{eff} is somewhat different. Thus if we ignore departures from LTE, we introduce a systematic error into estimates of T_{eff} . For main-sequence ($\log g = 4$) stars the errors can be neglected, whereas for $\log g = 3$ the errors are about 200°K and at $\log g = 2.5$ they are about 500°K ; these errors are small but systematic and

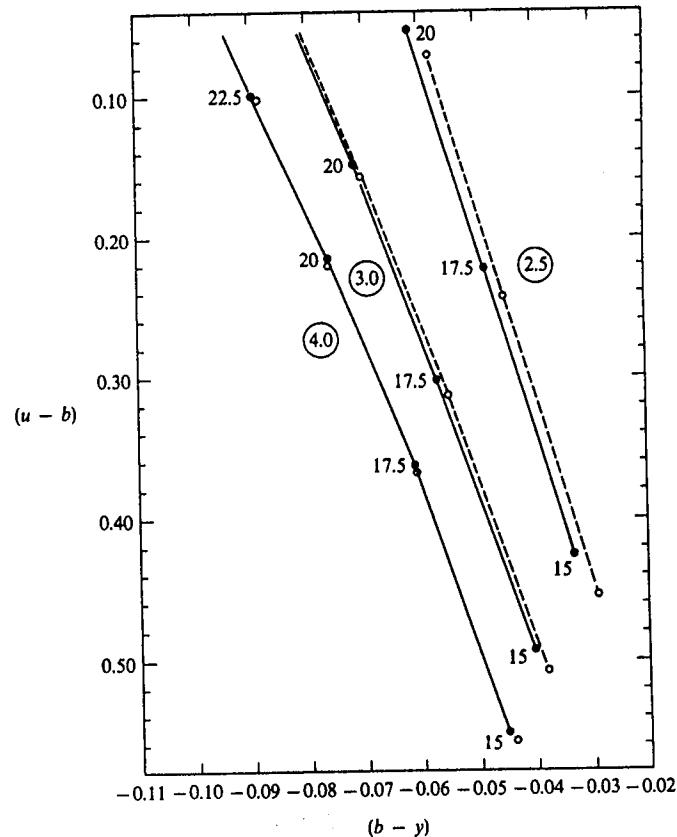


FIGURE 7-21
Theoretical Strömgren-system colors for LTE and non-LTE models.
Ordinate: $(u - b)$; abscissa: $(b - y)$. LTE values are represented by open circles and dashed curves; non-LTE values by solid dots and curves. Curves are labeled with $\log g$, and individual models are labeled with $T_{\text{eff}}/10^3$.

could have significant consequences in certain applications. The effects of departures from LTE upon the continuum jumps at the Lyman, Balmer, and Paschen edges (D_L , D_B , and D_P) are fairly substantial for B-stars. The Lyman jump is increased by non-LTE effects for most B-stars because the ground-state is overpopulated ($b_1 \gg 1$) and hence the flux for $\lambda < 912 \text{ \AA}$ is reduced; these changes are significant in making estimates of the far-ultraviolet energy output of these stars. Balmer jumps are generally decreased by non-LTE effects in B-stars because $b_2 < 1$ while $b_3 > 1$, so the opacity contrast across

the Balmer edge is reduced by departures from LTE. The Paschen jump is only slightly affected because both b_3 and $b_4 > 1$, and the opacity ratio stays about the same. In a plot of D_B vs $(b - y)$ (or some other flux ratio), the LTE and non-LTE models at a given gravity lie on the same curve, but slightly shifted (recall Figure 7-21). If one assigns values of T_{eff} to stars using such a diagram, the departures from LTE again imply systematic errors in estimates of T_{eff} ; these errors are negligible for main-sequence stars, are about 350°K at $\log g = 3$, and are about 500°K at $\log g = 2$.

From the point of view of obtaining direct observational evidence for departures from LTE in the continuum, the results mentioned above are not helpful because there is not a clear discrimination between the two cases. Strom and Kalkofen (611) pointed out that the parameter $\phi \equiv D_P/D_B$ provides a sensitive observational indicator of non-LTE effects, as may be seen in Figures 7-22 and 7-23. The LTE models predict values of ϕ of about 0.16 to 0.17, independent of gravity, whereas the non-LTE models predict much larger values of ϕ , increasing with decreasing gravity. This effect was found observationally (583); the supergiants have systematically larger values of ϕ than main-sequence stars with the same value of D_B . In the observational system there were some difficulties of calibration, so only a differential comparison was possible; these difficulties should be surmountable with the

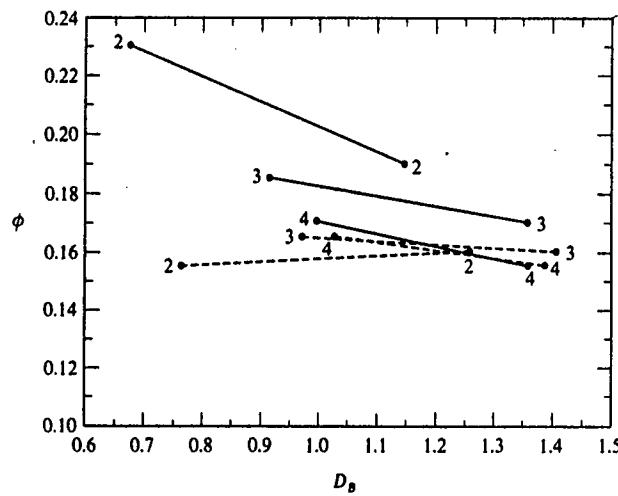


FIGURE 7-22
Theoretical continuum-jump parameters for LTE and non-LTE models of late B-type stars. Ordinate: $\phi \equiv D_P/D_B$; abscissa: Balmer jump D_B . Solid dots and curves: non-LTE values; open dots and dashed curves: LTE values. Curves are labeled with $\log g$.

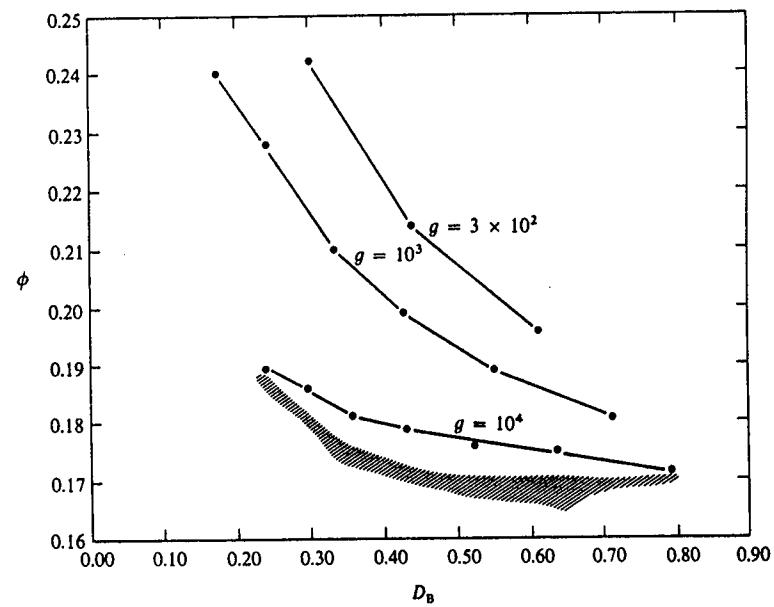


FIGURE 7-23
Same as Figure 7-22 for middle B-type stars. Shaded area contains predictions of LTE models at various values of g ; curves (labeled with surface-gravity g) show non-LTE values.

new Vega calibration. Further, departures from LTE have major effects on the ultraviolet continua of A-stars (587; 588) for which it is found that the ground-states of CI and SiI are strongly underpopulated, and the flux obtained from LTE calculations is too low by very large factors. Only when the non-LTE effects are taken into account is a satisfactory agreement with space observations obtained for these stars.

For O-stars ($T_{\text{eff}} \gtrsim 30,000^{\circ}\text{K}$) the effects of departures from LTE on visible continuum parameters and colors are much more important. For these stars, as described earlier, the ground state hydrogen becomes *underpopulated*, hence the flux below the Lyman limit *increases*, and the Lyman jump decreases. In contrast, at the 1227 Å ground-state edge of He II, the flux is decreased by non-LTE effects because $n = 1$ of He^+ is overpopulated [see (45; 432)]. These changes are of importance in estimating the energy output from O-stars into, say, nebulae or the interstellar medium. The most significant change in the visible region is that the Balmer jump predicted by the non-LTE models is about 0.07 mag larger (for $\log g = 4$) than that from LTE models, and shows but little variation with gravity, whereas the LTE

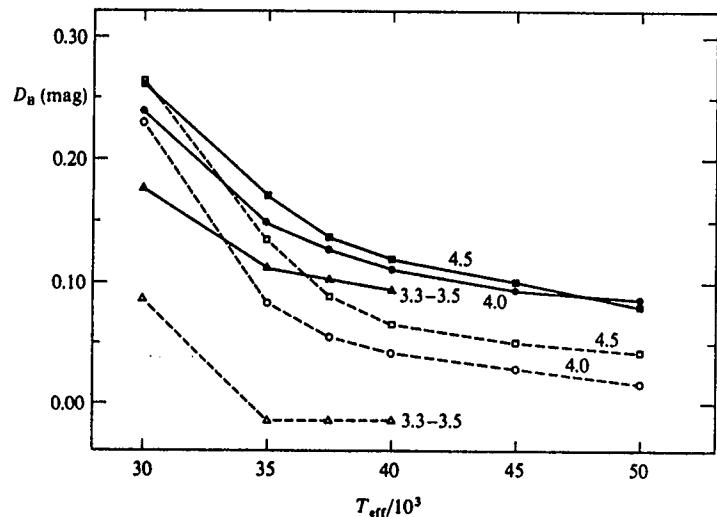


FIGURE 7-24
Theoretical Balmer jumps for O-stars. Ordinate: D_B in magnitudes; abscissa: $T_{\text{eff}}/10^3$. Solid symbols and curves: non-LTE models; open symbols and dashed curves: LTE models. Curves are labeled with $\log g$. From (45), by permission.

models predict emission edges at low gravities (see Figure 7-24). The absolute sizes of the non-LTE Balmer jumps are in excellent agreement with observations of O-stars if the Hayes calibration for Vega is adopted (as now seems correct). Further, the differential behavior of low-gravity versus high-gravity non-LTE Balmer jumps agrees with observation while the LTE results do not (405), and the same conclusion holds true for the differential behavior of O-star versus B0-star D_B 's (45); both of these results are independent of the calibration and support strongly the non-LTE calculations. The fact that the Balmer jump remains about constant in strength for the non-LTE models but weakens markedly in the LTE models implies very large differences in the computed colors [say ($u - b$) and ($b - y$)] of the two sets of models (516, 241). At the highest temperatures, the discrepancy in T_{eff} for a given value of ($u - b$) is some 15,000°K. Comparisons of T_{eff} derived from colors with those derived from helium-line strengths supports the non-LTE results. The comparison could be much refined yet by using a large body of data, and comparing the O-stars differentially with B-stars (thus eliminating calibration problems) [see also discussion in (516, 241)].

In summary, departures from LTE have significant effects on the visible continuum parameters of O-stars and B-type giants and supergiants, but are negligible for B-type main-sequence stars. Deviations from LTE produce

very large changes in the ultraviolet flux below the hydrogen Lyman edge, and in the resonance continua of certain light ions such as Cl and SiI. For later-type stars one must examine departures from LTE in the H⁻ ion. With present estimates of the relevant reaction rates (556; 188), the predicted effects (607) are negligibly small for G and K main-sequence stars and giants [see also (491; 492)]. Virtually nothing is known about the possible importance of departures from LTE in other types of stars (e.g., M giants or supergiants), and much work remains to be done to evaluate these effects. Finally, the problem of accounting for non-LTE effects in line-blanketing remains to be attacked.

TEMPERATURE STRUCTURE: THE CAYREL MECHANISM AND LINE EFFECTS

As we have seen in Chapter 3, the temperature in a grey atmosphere decreases uniformly outward to a limiting value $T_0/T_{\text{eff}} = 0.811$. Further, for a nongrey atmosphere that has a large opacity jump or strong lines, under the assumption of LTE the boundary temperature falls *below* the grey value (exceptions: scattering lines leave the boundary temperature unchanged and, in certain special cases, "new" absorbers can cause small temperature rises). In sum, the basic prediction for LTE radiative-equilibrium atmospheres is that the temperature distribution is a *monotone decreasing* function outward. Often the temperature structure exhibits plateaus where the then-dominant transition (e.g., the Balmer continuum) has become optically thin while other transitions (e.g., the Lyman continuum and Lyman lines) are completely opaque, followed by a series of drops as successively more opaque transitions become optically thin in turn.

For non-LTE atmospheres the situation is quite different, and typically the temperature distribution *passes through a minimum* and then shows a *rise outward*. The basic reason this occurs was pointed out by Cayrel (142; 283, 169), who called attention to the similarity between the physical conditions in the outer layers of a stellar atmosphere and those in a nebula surrounding a star. Suppose the energy balance is completely determined by absorption and emission from a single state (e.g., the sole bound state of H⁻, or the ground state of hydrogen). Then, in LTE, the condition of radiative equilibrium can be written

$$n_1^* \int_{v_0}^{\infty} \alpha_v B_v(T_0) dv = n_1^* \int_{v_0}^{\infty} \alpha_v J_v dv = n_1^* W \int_{v_0}^{\infty} \alpha_v B_v(T_R) dv \quad (7-160)$$

where the last equality introduces a parametric representation of J_v in terms of a dilution factor $W < 1$ and a radiation temperature T_R . As $W < 1$, it is clear that $T_0 < T_R$; in particular, if $W = \frac{1}{2}$ and $T_R = T_{\text{eff}}$, T_0 has essentially the grey value. The result just obtained follows because we have forced

LTE. But now suppose that we allow departures from LTE; then

$$n_i^* \int_{v_0}^{\infty} \alpha_v B_v(T_0) dv = b_1 n_i^* \int_{v_0}^{\infty} \alpha_v J_v dv = b_1 n_i^* W \int_{v_0}^{\infty} \alpha_v B_v(T_R) dv \quad (7-161)$$

In the limit of low density the departure coefficient is determined by the photoionization and recombination rates [cf. equation (5-95)], and

$$b_1 = \int_{v_0}^{\infty} (\alpha_v B_v/hv) dv / \int_{v_0}^{\infty} (\alpha_v J_v/hv) dv \quad (7-162)$$

Hence combining (7-161) and (7-162) we have

$$\begin{aligned} & \left[\int_{v_0}^{\infty} \alpha_v B_v(T_0) dv / \int_{v_0}^{\infty} \alpha_v B_v(T_0) v^{-1} dv \right] \\ &= \left[\int_{v_0}^{\infty} \alpha_v B_v(T_R) dv / \int_{v_0}^{\infty} \alpha_v B_v(T_R) v^{-1} dv \right] \quad (7-163) \end{aligned}$$

which shows that $T_0 = T_R$, independent of the value of W ! Thus we expect typically the temperature will decline from $T = T_{\text{eff}}$ near $\tau \approx 1$, to a value T_{min} , approximately equal to the grey-body boundary-temperature, and then rise to a value T_0 , with $T_{\text{min}} < T_0 \lesssim T_{\text{eff}}$. These results can also be viewed in terms of the "quantity" and "quality" of the radiation field (239). That is, in LTE it is the *energy density* (i.e., quantity) of the radiation that fixes T_0 ; in the non-LTE case, the energy density may be lower than its equilibrium value, but each photon has an energy characteristic of T_R and hence can still ionize the material and deposit an excess energy, *per ionization*, which is again characteristic of T_R [see also (240)].

For the sun, $T_{\text{eff}} = 5900^\circ\text{K}$, and Cayrel estimated $T_0 = T_R \approx 5600^\circ\text{K}$ compared to $T_{\text{min}} \approx 4800^\circ\text{K}$ (grey value). From a detailed calculation of a solar radiative equilibrium model, allowing for departures from LTE in the H^- ion, Feautrier obtained (211) $T_{\text{min}} \approx 4700^\circ\text{K}$ and $T_0 \approx 5200^\circ\text{K}$. Similar effects are found for O- and B-stars in which the main source of surface heating is the Lyman continuum. For example, in Figure 7-17 we saw that the LTE temperature structure of an atmosphere with $T_{\text{eff}} = 15,000^\circ\text{K}$, $\log g = 4$, had a plateau at $10,400^\circ\text{K}$, where the Lyman continuum was optically thick and all others were transparent, followed by a drop to $T_0 = 9400^\circ\text{K}$ when the Lyman continuum became transparent. In the non-LTE model, the temperature for $10^{-4} \lesssim \tau \lesssim 10^{-2}$ lies below the LTE value (because $b_2 < 1$ in the Balmer continuum and therefore the efficiency of heating is reduced). The non-LTE temperature distribution shows a minimum near $10,100^\circ\text{K}$ and then a rise outward to $T_0 \approx 10,350^\circ\text{K}$, which lies 1000°K above the LTE surface value. The agreement of this value of T_0 with the LTE Balmer plateau temperature is probably fortuitous. In fact, it is hard

to make an a priori estimate of an appropriate value of T_R in the Lyman continuum, for it, unlike H^- which is nearly grey, shows a sharp fall-off in opacity with increasing frequency, which implies that the surface radiation at high frequencies emerged from deeper, hotter layers and is characterized by larger values of T_R . The final result is established by some kind of average of T_R over frequency. Similar results are obtained at other effective temperatures (40; 211); detailed discussions of how particular transitions are affected may be found in the references.

Let us now consider the effects of lines on energy balance in the non-LTE case; here again it will repay the reader's effort to reread this material after Chapter 11 has been studied. A qualitative feeling for the results to be expected follows immediately from the form of the non-LTE source function for a collision-dominated line (cf. §11-2), namely

$$S_i = (1 - \epsilon_i) J_i + \epsilon_i B_v \quad (7-164)$$

where J denotes the average of J_v against the line-profile ϕ_v . By an argument (625) exactly analogous to that leading to equation (7-86), we find a result that differs only trivially from that obtained there, even though the line-source function has a noncoherent rather than coherent scattering term. We therefore reach the same conclusion as before: for LTE ($\epsilon = 1$) there is a *large* drop in boundary temperature, but when $\epsilon \ll 1$, the lines have practically no effect on the boundary temperature. The argument is verified by a detailed picket-fence calculation (448) that allows the continuum to adjust self-consistently. In this work the line-strength is assumed to remain unchanged, which would be valid, say, for the resonance lines of the dominant ion of a given species (e.g., Ca II in the solar atmosphere).

A very detailed calculation of non-LTE line-blanketing effects in the solar atmosphere (17) leads to the conclusion that a self-consistent model yields $T_{\text{min}} \approx 4330^\circ\text{K}$, and that the outward temperature rise driven by the Cayrel mechanism in the continuum is strongly resisted by the lines. In early-type stars the effects of lines upon the energy balance can be studied in non-LTE models constructed with the complete-linearization technique (references given above); several interesting results emerge. For example, in Figure 7-17 we see the effect of including $L\alpha$ (only) in addition to the Lyman continuum. In LTE the boundary temperature drops about 1600°K , from 9400°K to 7800°K ; but in the non-LTE case the boundary temperature had risen to $T_0 \approx 10,350^\circ\text{K}$ (owing to heating in the Lyman continuum), and inclusion of $L\alpha$ produces a temperature drop of only about one-third that obtained in LTE, to 9800°K . This result is not surprising, for $\epsilon \ll 1$. The final temperature structure is relatively complex in the non-LTE case.

An even more interesting example is shown in Figure 7-25, in which we see the importance of the coupling between lines and continua. Here the model hydrogen atom had three levels, and the formation of the first three

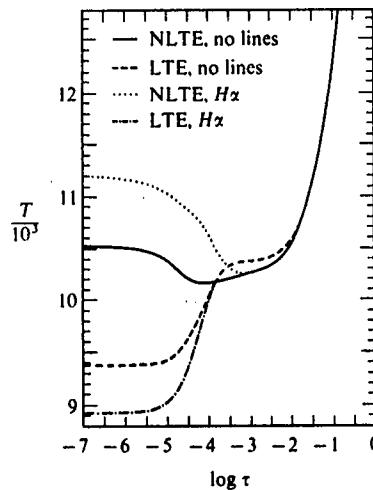


FIGURE 7-25
Temperature distribution for LTE and non-LTE models with $T_{\text{eff}} = 15,000^{\circ}\text{K}$ and $\log g = 4$. The atmosphere is composed of hydrogen, which is represented by a schematic model atom with three bound levels and continuum. This model atom accounts for the Lyman, Balmer, Paschen, and free-free continua, and the $H\alpha$ line. From (41), by permission.

continua, plus free-free, and the $H\alpha$ line was treated (41); the Lyman lines were omitted because they are formed in the very outermost layers, whereas $H\alpha$ is formed at about the same depth as the Lyman continuum and can interact with it. In LTE, $H\alpha$ decreases the boundary temperature from its continuum-only value of 9400°K to about 8900°K . In the non-LTE case, however, inclusion of $H\alpha$ raises the boundary temperature from about $10,500^{\circ}\text{K}$ (the continuum-only value) to $11,200^{\circ}\text{K}$; this is a decidedly non-classical result! The line itself makes a negative (i.e., cooling) contribution to the energy balance equation. But at the same time it provides an efficient channel for atoms to fall into the $n = 2$ level, where the radiation field produces strong heating; thus the direct cooling effect of the line is outweighed by indirect effects of the line on continuum energy balance, a possibility recognized by Cayrel (143). Addition of higher Balmer and Paschen-series lines (42) raises the temperature still further; the effect of $H\alpha$ alone is about equalled by inclusion of $H\beta$ and $P\alpha$, and yet-higher lines lead to only a small additional rise. A final example in Figure 7-26 shows the temperature structure in a model with $T_{\text{eff}} = 30,000^{\circ}\text{K}$, $\log g = 4$, (436) including the $L\alpha$, $L\beta$, Ly , $H\alpha$, $H\beta$, and $P\alpha$ lines. The temperature in the LTE model decreases uniformly outward; that in the continuum-only non-LTE model shows heating in the H Lyman and the He I and He II ground-state continua; the non-LTE model with lines shows enhanced heating from the Balmer lines followed by a drop produced by the Lyman lines.

The variety and complexity of the effects just described emphasizes the need for carrying out physically consistent analyses with great care. With the exception of the sun, for which a fairly detailed analysis exists, the prob-

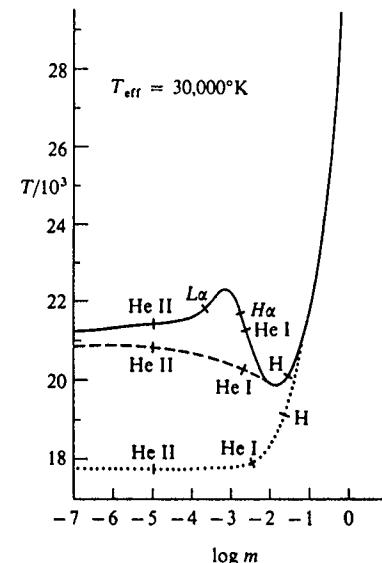


FIGURE 7-26
Temperature distributions in models with $T_{\text{eff}} = 30,000^{\circ}\text{K}$. Ordinate: $T/10^3$; abscissa: logarithm of column mass in gm cm^{-2} . Solid curve: non-LTE model including lines; dashed curve: non-LTE continuum-only model; dotted curve: LTE model. Cross-bars on curves mark optical depth unity in most opaque region of transition indicated (those with ion designations are ground-state continua). From (436), by permission.

lem of non-LTE line-blanketing in stellar atmospheres has barely been approached, and a large amount of work remains to be done.

7-6 Extended Atmospheres

For all of the models discussed thus far, it has been assumed that the atmosphere is stratified in plane-parallel layers; this is an excellent approximation when the density scale-height in the atmosphere is small compared to the radius of the star. However, many stars, in particular the supergiants and Wolf-Rayet stars, have *extended atmospheres* whose thicknesses are an appreciable fraction of a stellar radius; to a first approximation we suppose that these atmospheres are *spherically symmetric*. Atmospheric extension has important physical and observational implications. Thus stars with extended envelopes show a continuum energy distribution that has an anomalously low radiation temperature in comparison with the excitation temperature inferred from spectral lines. Equivalently, the energy distributions are "flatter", as a function of frequency, than those of main-sequence stars (which have compact, planar atmospheres) of the same spectral type, and show excess emission in the infrared and a deficiency in the ultraviolet. Almost always there are indications of rapid atmospheric expansion in stars with extended envelopes, so that one should consider dynamical models that include the effects of flow; we shall examine such models in Chapter 15, but for the

moment will deal with the more limited problem of solving the transfer equation in static extended envelopes.

In an extended, tenuous atmosphere, the radiation field at large distances from the underlying stellar disk becomes very dilute, and is confined primarily to a narrow solid angle (that subtended by the disk) around the radial direction. These facts imply that the temperature structure of the atmosphere must be quite different from that in a planar model, and that mathematical complications introduced by the angular peaking of the intensity will be encountered. The equation of transfer to be solved (cf. §2-3) is

$$\mu(\partial I_v/\partial r) + r^{-1}(1 - \mu^2)(\partial I_v/\partial \mu) = \eta_v - \chi_v I_v \quad (7-165)$$

with moments (cf. §2-4)

$$r^{-2}[\partial(r^2 H_v)/\partial r] = \eta_v - \chi_v J_v \quad (7-166)$$

and

$$(\partial K_v/\partial r) + r^{-1}(3K_v - J_v) = -\chi_v H_v \quad (7-167a)$$

or

$$[\partial(f_v J_v)/\partial r] + r^{-1}(3f_v - 1)J_v = -\chi_v H_v \quad (7-167b)$$

It is obvious that these equations are more complicated and more difficult to handle than their planar counterparts. For example, equation (7-165) is a *partial differential equation* involving explicitly two independent variables. Similarly the moment equations do not yield a simple form if, for example, we eliminate H_v between equations (7-166) and (7-167) (though a transformation will be described later that does allow a reduction of the two equations to a single combined moment equation of an attractive form). The moment equations could be dealt with, at least approximately, if we could relate K_v and J_v accurately, as is possible in the planar case. But now, even though $f_v \rightarrow \frac{1}{3}$ at great depth where the radiation field is homogeneous, $f_v \rightarrow 1$ near the surface [recall Exercise (1-12)], and a direct application of the Eddington approximation, which gives remarkably good results in the planar case, will not be even roughly correct near the surface. As was recognized by Mc Crea (413) and Chandrasekhar (148), it is possible to construct a consistent approximation scheme in a tenuous atmosphere that envelopes a parent star of radius r_* , by considering averages of the radiation field computed on the range $\mu_* \leq \mu \leq 1$ [where $\mu_* \equiv (1 - r_*^2/r^2)^{\frac{1}{2}}$] separately from those computed on the range $-1 \leq \mu \leq \mu_*$ [see also (403)]. But such a method is suitable only if r_* can be chosen unambiguously; for extended photospheres of appreciable density and nonnegligible optical depth this method breaks down [see, however, (635)]. It is thus not surprising that effective and general methods for treating transfer problems in spherical geometry have been slow to develop. A general and flexible numerical technique giving a direct solution of equations (7-165) through (7-167) will be presented later in this section. To treat the grey problem, an approximate

solution will be obtained merely by patching together asymptotic results for the two limiting regimes mentioned above; this provides a good starting point for estimation of extension effects on the emitted energy distribution.

SPHERICAL GREY ATMOSPHERES

The problem of grey spherical atmospheres in LTE and radiative equilibrium was analyzed approximately by Kosirev (360) and Chandrasekhar (148); numerical results of high precision have recently been obtained by Hummer and Rybicki (323). If we assume $\chi_v \equiv \chi$, and integrate equations (7-166) and (7-167) over all frequency, omitting the subscript v to denote integrated quantities, we obtain

$$r^{-2}[\partial(r^2 H)/\partial r] = 0 \quad (7-168)$$

$$\text{and} \quad [\partial(fJ)/\partial r] + r^{-1}(3f - 1)J = -\chi H \quad (7-169)$$

where in equation (7-168) we have demanded radiative equilibrium

$$\int_0^\infty \eta_v dv = \int_0^\infty \chi_v J_v dv = \chi J \quad (7-170)$$

Equation (7-168) yields an integral for the total flux,

$$r^2 H = H_0 = L/16\pi^2 \quad (7-171)$$

where L is the luminosity of the star. Define the optical depth, measured radially from an arbitrarily large outer radius R , to be

$$\tau(r) = \int_r^R \chi(r') dr' \quad (7-172)$$

Deep within the atmosphere (i.e., $r \ll R$, $\tau \gg 1$), we expect the radiation field to become isotropic and $f \rightarrow \frac{1}{3}$; in this limit equation (7-169) becomes

$$(\partial J/\partial r) = -3\chi H = -3\chi r^{-2} H_0 \quad (7-173)$$

which yields the integral

$$J(\tau) = H_0 \left(3 \int_0^\tau r^{-2} d\tau' + C \right) \quad (7-174)$$

If the usual Eddington-type boundary condition at $\tau = 0$ could be applied, then $J(0) = 2H(0) = 2H_0/R^2$ so that

$$J(\tau) = R^{-2} H_0 \left[3 \int_0^\tau (R^2/r^2) d\tau' + 2 \right] \quad (7-175)$$

a result obtained by Chandrasekhar (148). If, further, we impose the assumption of LTE, then $\eta_v = \chi_v B_v = \chi B_v$, and from equation (7-170) it follows that $J(\tau) = B(\tau) = \sigma T^4/\pi$, which assigns a temperature structure to the atmosphere. Equations (7-174) and (7-175) can be valid only at great depth. Near the surface, the free-flow regime occurs and $f \rightarrow 1$, in which case equation (7-169) becomes

$$[\partial(r^2 J)/\partial r] = -\chi r^2 H = -\chi H_0 \quad (7-176)$$

which yields

$$J(\tau) = r^{-2} H_0 (\tau + C) \quad (7-177)$$

In the limit $f = 1$, $J(0) = H(0) = H_0/R^2$, so equation (7-177) becomes

$$J(\tau) = r^{-2} H_0 (\tau + 1) \quad (7-178)$$

a result that is expected to be valid only when $\tau \ll 1$, $r \approx R$.

Considerable progress can now be made if we adopt a *power-law opacity* (i.e., $\chi = C_n r^{-n}$), as was done in the earliest work (360; 148). As noted by Kosirev, there is strong physical motivation for this choice in an expanding atmosphere, for the equation of continuity (see §15-1) demands that $\rho v r^2 = \text{constant}$, where ρ is the density, and v is the velocity of expansion; in the limit of very rapid expansion, when $v > v_{\text{escape}}$ (actually observed), the material moves at practically constant velocity, so that $\rho \sim r^{-2}$. The opacity of the material can be expected to vary as some power of the density (e.g., linearly for electron scattering or as the square for free-free) and hence as some power of $1/r$. Substituting the above expression for χ into equation (7-172) and adopting, for simplicity, $R = \infty$, we have

$$\tau(r) = C_n r^{-(n-1)/(n-1)} \quad (7-179)$$

Using equation (7-179), the limiting form for $J(\tau)$ (as $r \rightarrow 0$, $\tau \gg 1$) is

$$J \rightarrow [3(n-1)/(n+1)] H_0 r^{-2} \tau \quad (7-180)$$

while equation (7-177), valid for $\tau < 1$, remains unchanged. It is therefore attractive (387) to interpolate between these two extremes with an expression of the form

$$J(\tau) = \left(\frac{3H_0}{r^2} \right) \left(\frac{n-1}{n+1} \right) \left[\tau + \frac{1}{3} \left(\frac{n+1}{n-1} \right) \right] \quad (7-181)$$

Comparison with precise calculations demonstrates that equation (7-181) is quite accurate (323). Using the identity of J and B , and writing T_1 for the temperature at $\tau = 1$, we can rewrite equation (7-181) as

$$T(\tau) = T_1 \tau^{\frac{1}{2(n-1)}} \left[[\tau + (n+1)/3(n-1)] / [1 + (n+1)/3(n-1)] \right]^{\frac{1}{2}} \quad (7-182)$$

Equation (7-182) reveals an important characteristic difference between the planar and spherical cases, namely $T \rightarrow 0$ as $\tau \rightarrow 0$ in an extended atmosphere, rather than approaching a finite value. From this result we can see that the contribution from the outer cool layers, which occupy a large volume, will enhance the flux observed at longer wavelengths and lead to the distinctively flatter energy distribution mentioned above.

We may calculate the flux received by an observer at a great distance D from the center of the star by using the (p, z) coordinate system shown in Figure 7-27. The *impact parameter* p is the perpendicular distance of a ray from a parallel ray passing through the center of the star; z is the *distance along the ray*, measured from the plane through the center of the star perpendicular to the central ray. We shall take z to be positive towards the observer and formally place the observer at $z = \infty$ for purposes of calculating integrals. The coordinates (p, z) are related to polar coordinates (r, θ) as follows: $z = r \cos \theta$, $p = r \sin \theta$, and $r = (p^2 + z^2)^{\frac{1}{2}}$. If we choose a value of p , then the equation of transfer *along the ray* in the direction of increasing z is

$$(\partial I_v / \partial z) = \eta_v - \chi_v I_v \quad (7-183)$$

which follows from first principles. The formal solution of equation (7-183) for the intensity emerging at $z = \infty$ along the ray with impact parameter p may be written immediately as

$$I_v(p, \infty) = \int_{-\infty}^{\infty} B_v[T(p, z)] \exp[-\tau(p, z)] \chi(p, z) dz \quad (7-184)$$

where $\tau(p, z)$ is the optical depth measured from $z = \infty$, inward along the ray.

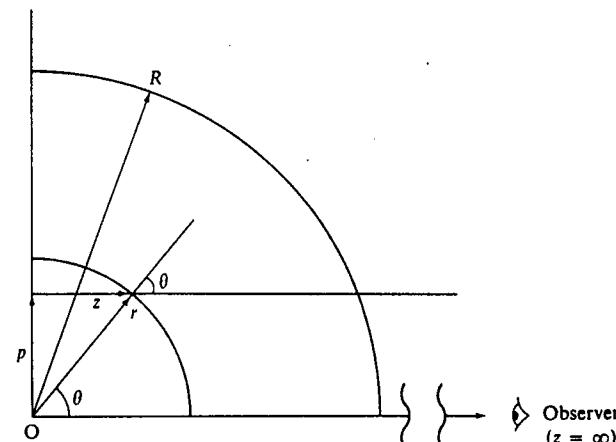


FIGURE 7-27
Coordinate systems for solution of transfer equation in spherical geometry.

Exercise 7-15: Calculate the operator $(\partial/\partial z)_p$, and show that equation (7-183) is exactly equivalent to equation (7-165). In the (p, z) coordinate system, the lines $p \equiv$ constant are the characteristic rays of the partial differential equation, which reduces to an ordinary differential equation along these particular curves.

The total flux received by the observer per unit receiver area is

$$f_v = 2\pi D^{-2} \int_0^\infty I_v(p, \infty) p \, dp \quad (7-185)$$

Following Kosirev (360) [see also (61, 165)] we change variables in equation (7-184) to $\theta = \cos^{-1}(z/r)$ and write $\tau(p, z) = \tau(p, \theta)$ which, using a power-law opacity, becomes $\tau(p, \theta) = C_n p^{-(n-1)} \psi_n(\theta)$ where

$$\psi_n(\theta) \equiv \int_0^\theta \sin^{n-2} \theta' \, d\theta' \quad (7-186)$$

Then

$$I_v(p, \infty) = C_n p^{-(n-1)} \int_0^\pi B_v[T(p, \theta)] \exp[-\tau(p, \theta)] \sin^{n-2} \theta \, d\theta \quad (7-187)$$

Substituting equation (7-187) into equation (7-185), and introducing successive transformations from $p = r \sin \theta$ to $\tau(r)$, the radial optical depth given by equation (7-179), we obtain

$$f_v = \pi(R_1/D)^2 \int_0^\infty B_v[T(\tau)] \tau^{-\left(\frac{2}{n-1}\right)} \Phi_n(\tau) \, d\tau \quad (7-188)$$

where R_1 is chosen as the radius at which $\tau = 1$, and

$$\Phi_n(\tau) \equiv 2 \int_0^\pi \exp[-(n-1)\tau \csc^{n-1} \theta] \psi_n(\theta) \sin \theta \, d\theta \quad (7-189)$$

Exercise 7-16: Derive equations (7-186), (7-188), and (7-189).

Using the temperature law given by equation (7-182) in equation (7-188), we may calculate the flux from a grey spherical atmosphere for a particular choice of T_1 , a characteristic temperature for the atmosphere, and the index n , which determines the degree of extension ($n \rightarrow \infty$ implies planar models; n small yields large extent). Observationally, color temperatures T_c are assigned, at a wavelength λ_c , by measurements of colors in a filter-system, or of the spectrophotometric gradient; both of these approaches, in effect, measure the slope of the continuum. Van Blerkom (61, 165) has calculated T_c at $\lambda_c = 5000 \text{ \AA}$ for various models with $T_1 = 50,000^\circ\text{K}$. He finds that, at $n = \infty$, $T_c/10^4 = 5$, and at $n = (10, 5, 3, \text{ and } 2)$, $T_c/10^4 = (4.3, 3.5, 2.2, \text{ and } 1.2)$ respectively, which shows that increasing atmospheric extent produces effects that simulate lower atmospheric temperatures. Indeed, a model with $T_1 = 50,000^\circ\text{K}$ and $n = 3$ has a flux distribution nearly identical to

one with $T_1 = 30,000^\circ\text{K}$ and $n = 5$. On the basis of these results we can understand why supergiants and WR stars have lower color temperatures than main-sequence stars of the same spectral type; it is also evident that atmospheric extension effects introduce ambiguities in the choice of a structural model for a star because of the trade-off between temperature and envelope-size in fitting the data. The frequency-variation of the flux from an extended grey atmosphere is shown in Figure 7-28, where it is compared to that from a Planck function at T_c , the color temperature of the flux at $\lambda_c = 5000 \text{ \AA}$, and to a Planck function at $T(\tau = \frac{2}{3})$ (the distribution that would emerge from a planar atmosphere). It is clear that the flux from an extended atmosphere shows a pronounced *ultraviolet deficiency* and an

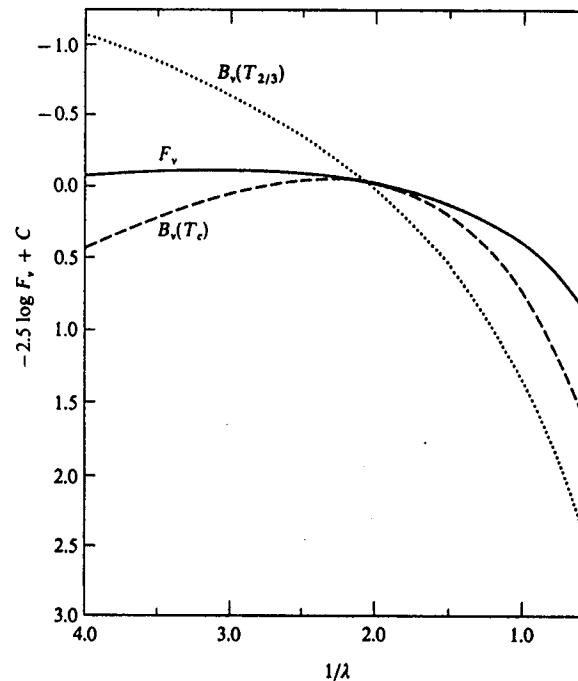


FIGURE 7-28

Solid curve: emergent flux F_v from spherical grey atmosphere with $T_1 = 5 \times 10^4 \text{ ^\circ K}$ and $n = 2$. Dashed curve: black-body curve at color temperature T_c that matches slope of F_v at $\lambda = 5000 \text{ \AA}$; note ultraviolet and infrared excesses of F_v relative to $B_v(T_c)$. Dotted curve: black-body curve at $T = T(\tau = 2/3)$, the value characteristic of a planar atmosphere. Abscissa: $1/\lambda$, where λ is in microns. From (61, 165), by permission.

infrared excess relative to that from a planar atmosphere of the same characteristic temperature. Relative to a Planck function at T_c , the temperature that would be assigned observationally, the flux distribution shows both an *infrared and ultraviolet excess*. These results emphasize the rather inhomogeneous nature of the radiation field from an extended atmosphere, arising from the greater variation of the temperature through the envelope. We shall consider more realistic nongrey and non-LTE models after developing a general method for solving the transfer equation.

SOLUTION OF THE TRANSFER EQUATION IN SPHERICAL GEOMETRY

A variety of methods have been employed to solve the transfer equation in spherical geometry. Both Kosirev (360) and Chandrasekhar (148) used variants of the Eddington approximation, which, however, is not accurate near the surface. Subsequent work (154; 679; 680) approximated the angular behavior of the radiation by an expansion in spherical harmonics, but only to second order. As demonstrated convincingly by Chapman (163), these methods cannot be adequate because the radiation becomes very sharply peaked in the outward direction ($f_k \rightarrow 1$) at the surface [see (163, Fig. 1) or (323, Fig. 6)]. Methods that employ direct differencing of both the angular and spatial derivatives in equation (7-165), [related to the S_n -method (127) of reactor physics] have been suggested (130; 255; 507), but these fail when the discrete spherical shells become optically thick, and hence must be supplemented by special techniques to be useful for stellar atmospheres work.

The differential-equation technique to be described below, which is general, stable, and efficient, uses a Feautrier solution along individual impact parameters (tangent to discrete shells) to generate the angular information required to evaluate variable Eddington factors (323) from an estimated source function. Then, with known Eddington factors, it employs a Feautrier scheme to solve a combined moment equation obtained by using the elegant transformation introduced by Auer (32). Alternatively, a direct solution can be obtained with a Rybicki-type method, if the scattering integral in the source function is frequency-independent (442); an equivalent integral-equation method has also been developed (558).

Consider first the moment equations; regard all variables as functions of r and v , and introduce the radial optical depth scale $d\tau_v = -\chi_v dr$. Then the moment equations are

$$\partial(r^2 H_v)/\partial\tau_v = r^2(J_v - S_v) \quad (7-190)$$

and

$$[\partial(f_v J_v)/\partial\tau_v] - (3f_v - 1)J_v/(\chi_v r) = H_v \quad (7-191)$$

where S_v is assumed to have a general form

$$S_v = \alpha_v \int R(r; v', v) J_{v'} dv' + \beta_v \quad (7-192)$$

For a pure continuum problem, we would have only a coherent scattering term and S_v would be simpler. There are two essential difficulties with equations (7-190) and (7-191). (a) Direct elimination of H does not yield a simple equation, but rather a complicated one involving both first and second derivatives. (b) The term in equation (7-191) in $(\chi_v r)^{-1}$ tends to diverge strongly at the surface (recall χ_v is opacity per unit volume and hence varies with the particle number density over several orders of magnitude). This term destabilizes the system. Both of these difficulties can be eliminated entirely by the introduction (32) of a *sphericality factor* q_v , defined by

$$\ln(r^2 q_v) = \int_{r_c}^r [(3f_v - 1)/(r' f_v)] dr' + \ln r_c^2 \quad (7-193)$$

where r_c is a "core radius" corresponding to the deepest point in the atmosphere considered in the solution. It is obvious that q_v is known if f_v is known. The factor q_v allows equation (7-191) to be rewritten as

$$\partial(f_v q_v r^2 J_v)/\partial\tau_v = q_v r^2 H_v \quad (7-194)$$

which, when substituted into equation (7-190), yields a *combined moment equation*

$$\frac{\partial}{\partial\tau_v} \left[\frac{1}{q_v} \frac{\partial(f_v q_v r^2 J_v)}{\partial\tau_v} \right] = r^2(J_v - S_v) \quad (7-195)$$

or, introducing a new variable $dX_v = -q_v \chi_v dr = q_v d\tau_v$,

$$\partial^2(f_v q_v r^2 J_v)/\partial X_v^2 = q_v^{-1} r^2 (J_v - S_v) \quad (7-196)$$

Exercise 7-17: Verify that equation (7-193) allows the reduction of equations (7-190) and (7-191) to (7-195).

To obtain an upper boundary condition we define

$$h_v \equiv \int_0^1 I(R, \mu, v) \mu d\mu / \int_0^1 I(R, \mu, v) d\mu \quad (7-197)$$

so that from (7-194)

$$[\partial(f_v q_v r^2 J_v)/\partial X_v]_{r=R} = h_v (r^2 J_v)_{r=R} \quad (7-198)$$

while at the lower boundary we apply a planar diffusion approximation to write

$$H_v(r_c) = \frac{1}{3} (\chi_v^{-1} |\partial B_v / \partial r|)_{r_c} \quad (7-199)$$

and fix the gradient by demanding that the integral of $H_v(r_c)$ over all frequencies equal the correct integrated flux $H_c = L/(16\pi^2 r_c^2)$. Then

$$[\partial(f_v q_v r^2 J_v) / \partial X_v]_{r=r_c} = r_c^2 H_c \left[\chi_v^{-1} (\partial B_v / \partial T) \left/ \int_0^\infty \chi_v^{-1} (\partial B_v / \partial T) dv \right. \right]_{r=r_c} \quad (7-200)$$

The diffusion approximation, and hence equations (7-199) and (7-200), will be valid when the photon mean free path $\chi_v^{-1} \ll \epsilon R$ where ϵ is some small number; this criterion can always be met by choosing r_c sufficiently deep in the atmosphere. For other physical situations (e.g., in a nebula), alternative inner boundary conditions can be posed (374).

With the introduction of a discrete radius mesh $\{r_d\}$, ($d = 1, \dots, D$), where $R = r_1 > r_2 > \dots > r_D = r_c$, and a frequency mesh $\{v_n\}$, ($n = 1, \dots, N$), we may replace derivatives with finite differences [perhaps using splines (374; 442) or Hermite formulae (34)] and the frequency integral (if any) in the source function with a quadrature formula. Equations (7-196), (7-198), and (7-200) are then of the standard tridiagonal form of equation (6-31) and can be solved with the usual Feautrier elimination scheme; the computing time scales as $T_M = c DN^3$, which is particularly economical if the source function is purely thermal or has only a coherent scattering term ($N = 1$). In the calculation just described, the scattering integral in the source function appears explicitly, hence the correct global thermalization properties of the solution are obtained.

To carry out the computations described above, we require knowledge of the Eddington factors f_i ; these can be found if we know the angular behavior of the radiation field at each depth. To obtain the required information we perform, at each frequency, a ray-by-ray solution along a grid of impact parameters $\{p_i\}$, chosen to be tangent to each discrete radial shell, augmented by an additional set of C impact parameters, which are chosen to intersect the core, and which include the central ray. The geometry of the situation is shown in Figure 7-29. The impact parameters are labelled with an index i , ($i = 1, \dots, I$), where $I = D + C$; p_1 denotes the central ray; p_c the last ray *inside* the core (i.e., $p_c < r_c$); $p_{c+1} = r_c$; and $p_I = R$. Each ray p_i intersects all shells with $r_d \geq p_i$, and these intersections define a mesh of z -points $\{z_{di}\}$, ($d = 1, \dots, D_i$). Here $D_i = D + C + 1 - i$ for $i > C$, $D_i = D$ for $i \leq C$, and $z_{di} = (r_d^2 - p_i^2)^{\frac{1}{2}}$. It is seen from the figure that the ray p_i intersects the radial shell r_d at an angle whose cosine is

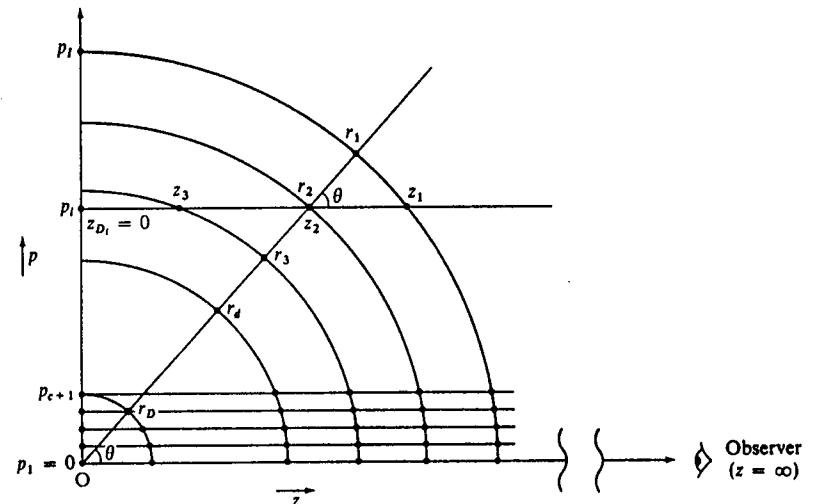


FIGURE 7-29

Discrete (p, z) mesh used in solution of spherical transfer equation. The impact parameters $\{p_i\}$ are chosen to be parallel to the central ray, and tangent to spherical shells chosen to describe the depth-variation of physical properties of the envelope. The intersections of the rays with the radial shells define a z -mesh along each ray.

$\mu_{di} \equiv \mu(r_d, p_i) = (r_d^2 - p_i^2)^{\frac{1}{2}}/r_d = z_{di}/r_d$. Hence if we first construct the solution along all rays $\{p_i\}$, and choose a particular value r_d , then knowledge of the variation of $I_v(z_{di}, p_i)$ ($i = 1, \dots, I_d$) is equivalent to knowledge of the μ -variation of $I_v(r_d, \mu)$ on the mesh $\{\mu_{di}\}$, ($i = 1, \dots, I_d$), which spans the interval $1 \geq \mu \geq 0$. Here $I_d \equiv I + 1 - d$. It is therefore clear that the ray-by-ray solution described above will allow determination of the requisite Eddington factors. The astute geometrical trick employed here to synthesize the angular information from ray solutions is actually nontrivial and makes direct use of the symmetry of the problem, which allows us to treat all points on a given shell as equivalent; without strict spherical symmetry, the problem is much more complex.

Consider now the ray specified by p_i . The transfer equation along the ray is

$$\pm [\partial I^\pm(z, p_i, v) / \partial z] = \eta(r, v) - \chi(r, v) I^\pm(z, p_i, v) \quad (7-201)$$

where the + and - signs refer, respectively, to radiation flowing toward and away from the external observer, and we have written r as the space variable in η and χ with the understanding that $r \equiv r(z, p_i) = (p_i^2 + z^2)^{\frac{1}{2}}$. Defining the optical depth along the ray $d\tau(z, p_i, v) \equiv -\chi(r, v) dz$, setting $S(r, v) \equiv \eta(r, v)/\chi(r, v)$ (assumed to be known values), and introducing the

mean-intensity-like and flux-like variables

$$u(z, p_i, v) \equiv \frac{1}{2} [I^+(z, p_i, v) + I^-(z, p_i, v)] \quad (7-202)$$

and $v(z, p_i, v) \equiv \frac{1}{2} [I^+(z, p_i, v) - I^-(z, p_i, v)] \quad (7-203)$

we obtain the second-order system

$$[\partial^2 u(z, p_i, v)/\partial z^2(z, p_i, v)] = u(z, p_i, v) - S[r(z, p_i), v] \quad (7-204)$$

with an upper boundary condition

$$\partial u(z, p_i, v)/\partial z(z_{\max}, p_i, v)|_{z_{\max}} = u(z_{\max}, p_i, v) \quad (7-205)$$

where $z_{\max} = (R^2 - p_i^2)^{\frac{1}{2}}$. The inner boundary condition depends upon whether the ray intersects the core, $r_{D_i} = r_c$, or misses the core and intersects the central plane at $z = 0$. In the former case we apply the diffusion approximation as was done to obtain equation (7-200); in the latter, symmetry considerations show that $v(0, p_i, v) \equiv 0$, hence

$$\partial u(z, p_i, v)/\partial z(z=0) = 0 \quad (7-206)$$

Equations (7-204) through (7-206), when written as difference equations, yield (with known S) a single tridiagonal system of the standard Feautrier form, and can be solved by the usual algorithm. The computing time for N frequencies, C core-rays, and D radial depth-points scales as $T_R = cN[D \cdot C + \sum D_i] \approx c'ND^2$ for $D \gg C$. Having calculated the complete solution $u_{din} \equiv u(z_d, p_i, v_n)$ we construct the moments, as described above,

$$J_{dn} = \sum_{i=1}^{I_d} w_{di}^{(0)} u_{din} \quad (7-207)$$

and $K_{dn} = \sum_{i=1}^{I_d} w_{di}^{(2)} u_{din} \quad (7-208)$

and thus the Eddington factor $f_{dn} \equiv K_{dn}/J_{dn}$. Here the w 's are appropriate quadrature weights, obtained analytically by integration of moments of a piecewise polynomial representation of $u(r_d, \mu)$ on the mesh $\{\mu_i\}$, generated by the intersection of the rays $\{p_i\}$ with the radial shell r_d . Using the new Eddington factors, the moment equations are re-solved and the process is iterated to convergence; experience shows the convergence to be very rapid.

Exercise 7-18: (a) Write difference approximations for equations (7-204) through (7-206). (b) Derive a second-order accurate lower boundary condition by ex-

tending the difference equation for equation (7-204) beyond the central plane $z_{D_i} = 0$, and using symmetry of u_{din} about this plane.

In the case that the source function contains a single scattering integral involving J (or \bar{J} in a line), rather than a frequency-dependent scattering term involving partial redistribution, we may avoid the iteration procedure between ray equations and moment equations, and develop a *direct* solution by a Rybicki-type scheme (442). Along each ray, for each frequency, we have a tridiagonal system of the form

$$T_{in} \mathbf{u}_{in} = \mathbf{U}_{in} \mathbf{J} + \mathbf{W}_{in} \quad \begin{aligned} (i &= 1, \dots, I) \\ (n &= 1, \dots, N) \end{aligned} \quad (7-209)$$

where \mathbf{J} describes the depth-variation of J [cf. equation (6-46)] on the range ($d = 1, \dots, D_i$), and \mathbf{u}_{in} the variation of $u(z, p_i, v_n)$ along the ray. This system can be solved for $\mathbf{u}_{in} = \mathbf{C}_{in} \mathbf{J} + \mathbf{D}_{in}$, and this solution can be substituted into the equation defining \mathbf{J} [cf. equation (7-207)] to yield a final system for \mathbf{J} ; the computing time for this solution scales as $T_D = cN D^3$ for $D \gg C$. Details of this procedure, which is general, stable, and economical, can be found in the reference cited.

EXTENDED MODELS FOR EARLY-TYPE STARS

Nongrey spherical model atmospheres in LTE have been constructed for the central stars of planetary nebulae (130; 131; 376) and for O and B supergiants (136; 325; 376; 441; 442; 559; 516, 241), using a variety of techniques. Non-LTE models for both classes of objects are given in (376; 441; 442; 516, 241). All of these models assume hydrostatic equilibrium, and the extension effects are produced by a near-cancellation of gravity by radiation forces on the material. In actuality, there is strong evidence that atmospheric extension is almost always associated with large-scale expansion, and static models can be expected to yield, at best, only qualitative information; we shall discuss *dynamical* models, which are more difficult to construct, in §15-4.

In an extended atmosphere, it is necessary to account for the variation of the gravitational force with radius, and if M is the mass of the star, the equation of hydrostatic equilibrium becomes

$$(dp_0/dr) = -(\rho G M / r^2) + (4\pi/c) \int_0^\infty \chi_v H_v dv \quad (7-210)$$

Introducing the Rosseland mean opacity χ_R , the integrated flux $H = L/(16\pi^2 r^2)$, and the parameter

$$\gamma \equiv (\chi_R H)^{-1} \int_0^\infty \chi_v H_v dv \quad (7-211)$$

$$\text{we have } [dp_\theta/d(1/r)] = \rho[G\mathcal{M} - (\gamma L\chi_R/4\pi c\rho)] \quad (7-212)$$

As was true for planar atmospheres, radiation forces lead to an increase in the scale-height (and hence the extent) of the atmosphere. The parameter $\Gamma = \gamma\chi_R L/(4\pi c G\mathcal{M}\rho)$ measures the ratio of radiation to gravity forces. In the limit of a pure electron-scattering opacity we find an upper bound on the critical ratio for (L/\mathcal{M}) at which radiation forces just balance gravity forces, namely $(L/\mathcal{M})_{\text{crit}} = 3.8 \times 10^4 (L_\odot/\mathcal{M}_\odot)$. (cf. Exercise 7-1). We shall denote the value of Γ obtained from pure electron scattering as Γ_e . For an extended envelope, the meaning of the stellar "radius" (and hence of the effective temperature) becomes ambiguous; typically one uses the value r_3 , at which $\tau_R = \frac{3}{2}$, as a characteristic radius.

The effects of radiation forces are quite important in the planetary nebula stars, as estimates of L and \mathcal{M} for these objects (279) lead immediately to values of Γ on the range $0.8 < \Gamma_e < 0.93$. Further, it can be shown (442) that the absolute thickness of the atmosphere, Δr , is proportional $g_{\text{eff}}^{-1} = [(1 - \Gamma)G\mathcal{M}/R^2]^{-1}$, which means that for a given g_{eff} , the relative thickness of the atmosphere, $\Delta r/R$, will be larger for stars of small radius, such as the planetary nebula stars. Thus several of the models in (130; 131) have atmospheric thicknesses comparable to, or even larger than, the radius of the "core" star (depths below $\tau = 10$!). The effect of extension in these stars greatly reduces the color temperature relative to the effective temperature; at $\lambda_c = 5000 \text{ \AA}$, T_c/T_{eff} for some of the models is as small as 0.3. The complete energy distribution shows the characteristic flattening described above for grey models [indeed in the visible, the nongrey energy distribution can be fitted quite accurately by grey models with $n = 3$ to 4 (61, 165)]. The flatter energy distributions predicted by these models strongly resembles those of WR stars (though the model parameters L , \mathcal{M} , and R are not really appropriate for such stars).

For O-type supergiants and Of-stars, it is clear, from the failure of planar models to match the energy distribution [see (376)] and to produce the observed emission lines, that the atmosphere must be extended. On the other hand, stellar evolution tracks at plausible masses all yield values of $\Gamma_e < 0.5$, which, taking into account the large radii of these stars, implies insignificant atmospheric extension (136). But absorption in spectrum lines can greatly increase the total force of the radiation on the material (102, 404) and indeed it is precisely this mechanism that is thought to drive the winds from O-stars (see §15-4). In several calculations (376; 441; 442; 516, 241), attempts were made to simulate the effects of enhanced radiation forces by adopting ad hoc values for the radiation-force multiplier γ in equation (7-212), chosen so as to raise the maximum value of Γ to values approaching unity (the most extreme case considered had $\Gamma_{\text{max}} \approx 0.995$). Although such models become very difficult to compute because of numerical instabilities, which directly reflect the physical instability of the atmosphere, a number of solutions were

obtained using a complete linearization procedure, both assuming LTE and taking into account departures from LTE. It should be stressed that, for such extreme values of Γ , hydrostatic equilibrium is very unlikely to be possible, and dynamical models are necessary; models such as these merely demonstrate the effects of a greatly increased scale-height for the envelope. Extension effects produce major changes in observable parameters such as colors, see Figure 7-30, which show a strong reddening with increasing atmospheric size. The colors for LTE and non-LTE models differ markedly because in LTE a spurious emission edge occurs at the Balmer jump (caused by the

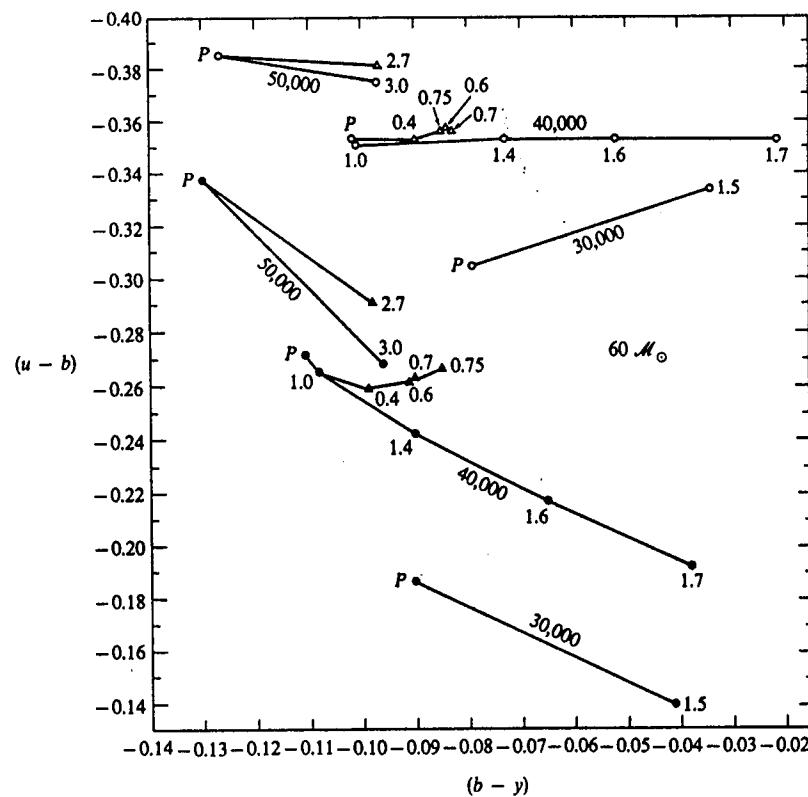


FIGURE 7-30
Strömgren-system $(u - b)$ and $(b - y)$ colors for models with $\mathcal{M} = 60 \mathcal{M}_\odot$. Open symbols: LTE models; filled symbols: non-LTE models. Models with equal values of T_{eff} are joined on curves. Curves are labeled with parameters describing radiation-force multiplier γ . Circles: $\gamma = \gamma_1$, labels give value of γ_1 ; triangles: $\gamma = 1 + \gamma_2 \exp(-\tau_R)$, labels give value of γ_2 . From (376), by permission.

Schuster mechanism—see §10-2); detailed discussion of the overall properties of the energy distributions can be found in the references cited above.

7-7 Semiempirical Solar Models

All of the discussion presented thus far in this book has concerned the development of methods that allow the analysis of a stellar spectrum, and the inference of physical properties of the atmosphere, by a model-fitting technique. Our information about stars, which we perceive as mere points of light, is quite limited in quantity, quality, and scope, primarily because of the low intensity of the radiation received. But for one star, the sun, we have unparalleled opportunities to obtain information at very high spatial, temporal, and spectral resolution, over an enormous range of energies, from X-ray, ultraviolet, visible, infrared, and radio wavelengths. While stars are unresolved, structures of the order of 150 km in size can be distinguished in the solar atmosphere, and a wide variety of small-scale features can be seen. Further, measurements of velocity fields and magnetic fields can be made. The sun thus provides a unique testing ground for our theoretical methods, and, because so much information can be obtained by semiempirical analyses that rely only weakly upon theory, it offers guidance in seeing how to extend the theory in those places where it is oversimplified. It is fair to say that most (though not all) major advances in stellar atmospheres theory were motivated by attempts to understand solar phenomena; time and again, the refined confrontations possible between solar observations and the theory have forced radical changes in our interpretive picture. And many enigmas remain, so we may expect much further development yet.

There is no hope of summarizing adequately, in a few pages, the vast amount of information known about the solar atmosphere, and the reader should consult the many excellent references available, particularly (244; 694; 11, Chaps. 9 and 10) for general information, (628; 19; 20) for detailed information about the chromosphere, and (94; 20) for material about the corona. The goal of the discussion below will be only to present a brief summary of some of the primary structural properties of the solar atmosphere, in order to provide background and orientation for the development of more realistic pictures of stars in general.

In terms of basic morphology, the solar atmosphere can be divided into four major parts: (a) the *photosphere*, the opaque disk seen visually; (b) the *chromosphere*, a region extending some 2500 km above the limb, showing the characteristic emission spectrum of hydrogen; (c) the *corona*, a tenuous, faint outer envelope first seen at eclipses; and (d) the *wind*, a supersonically expanding region that streams past the earth. The overall temperature structure of the atmosphere is shown in Figure 7-31. In the photosphere the temperature decreases outward to a minimum of about 4200°K. At the relevant

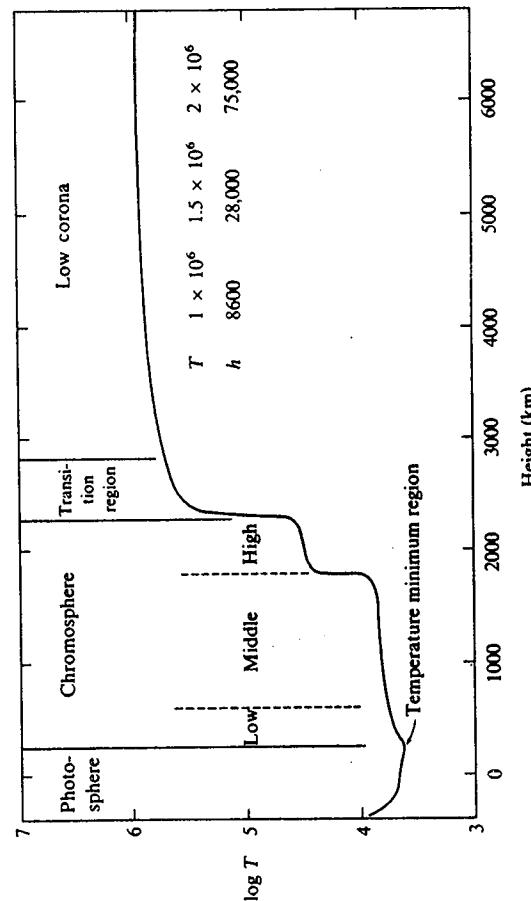


FIGURE 7-31
Temperature as a function of height in the solar atmosphere. The zero-point of the height-scale in this plot refers to unit tangential optical depth at the limb (radial optical depth $t_{5000} \approx 3.7 \times 10^{-3}$), and lies at approximately +340 km on the scale of Figure 7-32 and Table 7-1. From (20) by permission.

temperature ($T \sim 5 \times 10^3$ °K) and gravity ($g \sim 3 \times 10^4$) the pressure scale-height is about 120 km. This length subtends an angle of 0.15", which lies well below the typical limit of resolution set by seeing effects in the earth's atmosphere; as a result, the sun appears to have a crisply defined edge, which occurs at height $h = 0$ km on the scale of Figure 7-31. In the low and middle chromosphere the temperature rises to about 8000°K and then to a plateau of around 30,000°K. There is then a very thin *transition region* in which the temperature suddenly rises to coronal values of about 1.5×10^6 °K. In the corona the scale-height is about 50,000 km, which is a significant fraction of a solar radius, so the corona extends to large distances from the sun; indeed, even if the corona were static, it would envelope the earth with a substantial density of particles.

The temperature structure described above has been inferred mainly from analysis of the solar spectrum, which is strongly affected by the nature of the temperature distribution. The features in each spectral region arise from some characteristic range of height (see Figure 7-32) and, since it has become possible to survey a wide range of wavelengths from space vehicles, our knowledge of the temperature structure has been greatly improved. In the spectral range from $1685 \text{ \AA} \lesssim \lambda \lesssim 350 \mu$ the continuum originates in the photosphere, and the lines have unit optical depth at heights that extend from the photosphere (in the line wings) into the chromosphere (at line center). Observed on the disk, lines with wavelengths $\lambda \gtrsim 1900 \text{ \AA}$ are dominantly in absorption, and a transition to emission lines occurs in the range $1700 \text{ \AA} \lesssim \lambda \leq 1900 \text{ \AA}$. The continuum on the range $1525 \text{ \AA} \lesssim \lambda \lesssim 1685 \text{ \AA}$ comes from the photosphere-chromosphere transition region, and for $\lambda < 1525 \text{ \AA}$ is dominantly chromospheric. On the range $504 \text{ \AA} \lesssim \lambda \lesssim 912 \text{ \AA}$ the Lyman continuum is in emission; for $\lambda \lesssim 504 \text{ \AA}$ the emission is from the continuum of He I. Chromospheric emission lines extend at least down to 288 \AA , through the resonance series of He II. Coronal emission lines begin to appear prominently at about 800 \AA , and extend down into the X-ray region. In the infrared, one sees most deeply into the atmosphere near 1.6μ , the minimum in the H^- opacity. At longer wavelengths, the temperature minimum region is encountered near 300μ ; and at radio wavelengths of 100 cm and beyond, the continuum emission is entirely coronal. (The corona is completely transparent in the continuum at centimeter wavelengths and shortward.) Off the disk, the chromosphere and corona can be seen in strong emission lines, by using special instrumentation, or in special circumstances such as eclipses, from which an enormous wealth of data has been derived.

A variety of analytical techniques have been used to make estimates of physical properties of the solar atmosphere from the data described above. A very powerful tool is the analysis of *limb-darkening* data. As we saw in §3-3, the Eddington-Barbier relation implies that the dominant contribution to $I_r(0, \mu)$ comes from $S_r(\tau_r \approx \mu)$; hence by scanning from center to limb,

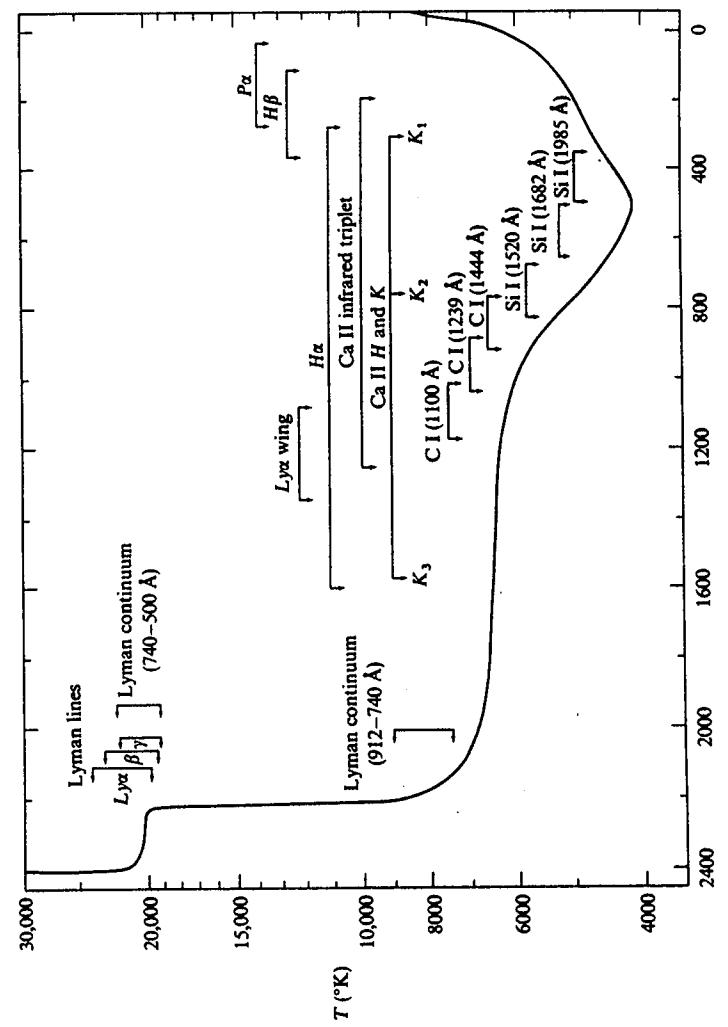


FIGURE 7-32
Temperature as a function of height in the solar atmosphere. Unit optical depth for various lines and continua (over their entire profiles) are indicated by lines with arrows. From (645), by permission.

$0 \leq \mu \leq 1$, we sample depths $0 \lesssim \tau_v \lesssim 1$. Suppose we assume that the source function is the Planck function, and that we can write $B_v(\tau_v)/I_v(0, 1) \equiv b_v(\tau_v)$. Then

$$\Phi_v(\mu) = [I_v(0, \mu)/I_v(0, 1)] = \int_0^\infty b_v(\tau_v) \exp(-\tau_v/\mu) d\tau_v/\mu \quad (7-213)$$

If we now assume that b_v can be represented by some analytical formula, Φ_v can be computed in terms of the expansion coefficients of the formula. By fitting this calculated function to the observations of Φ_v , one finds the coefficients and hence b_v . For example, if we take

$$b_v = \sum_k a_k \tau_v^k \quad (7-214)$$

then we find

$$\Phi_v(\mu) = \sum_k a_k k! \mu^k \quad (7-215)$$

Thus a fit of equation (7-215) to the data yields the a_k 's, hence b_v and, ultimately, if we know $I_v(0, 1)$ we obtain $B_v(\tau_v)$, and hence $T(\tau_v)$. A very large body of limb-darkening data exists, extending from ultraviolet through radio wavelengths.

The procedure outlined above encounters many practical difficulties. Thus even large-amplitude fluctuations in the physical properties do not affect the observations if they occur in regions whose line-of-sight optical thickness is $\ll 1$; we thus remain in ignorance of any such inhomogeneities. Similar remarks apply to horizontal structures that lie below the limit of resolution set by the observing procedure. Moreover, we clearly obtain but little information from depths with $\tau_v > 1$, while a limit is also set at small depths by seeing effects at the limb; a typical resolution of $1''$ implies a blur of about 0.001 solar radius, which corresponds to $\mu \approx 0.05$. Thus at any given wavelength we are confined to $0.05 \lesssim \tau_v \lesssim 1$. Further, with data of limited quality the number of expansion coefficients that yield real information in equation (7-214) or (7-215) is restricted. For example, for data with an accuracy of $\pm 1\%$, at most three coefficients can be obtained (98); attempts to use more coefficients yield results that fit the observed values of $\Phi_v(\mu)$ but show wild oscillations in $b_v(\tau_v)$. Numerical methods (such as the *Prony algorithm*) exist that automatically limit the coefficients to only those justified by the accuracy of the data (669). One may choose a variety of fitting functions in equation (7-214), each of which fits the observed $\Phi_v(\mu)$ equally well; in general these yield different values of $b_v(\tau_v)$ and hence of $T(\tau_v)$. Among three-term solutions, discrepancies in the temperature of the order of $\pm 200^\circ\text{K}$ are found; these can be important in certain contexts. For example, in the solar atmosphere the dominant ionization state of iron atoms is Fe^+ . Using the Saha equation it is easy to show that an error of $\pm 200^\circ\text{K}$ introduces an uncertainty in the

occupation numbers of low-lying states of Fe I (7.9 eV below the ionization limit) of about $\Delta \log n \approx \pm 0.25$, or about $\pm 60\%$; thus even assuming that every other step in the analysis of the spectrum were perfectly accurate, significant errors are introduced into estimates of element abundances from uncertainties in the model alone.

The range of depths available to analysis using limb-darkening data may be extended by using different wavelengths. One then needs to map the curves of $T(\tau_v)$ from one frequency to another, which can be done if we know the frequency variation of the absorption coefficient—i.e., $\chi(v)/\chi(v_{\text{std}})$. Alternatively, we can demand that the different $T(\tau_v)$ curves correspond to a unique $T(h)$ variation, and turn the procedure around to infer the frequency profile of the absorption coefficient. Studies of this kind (e.g., 517) have shown that the empirically-determined absorption coefficient is consistent with the hypothesis that H^- is the primary opacity source (though additional sources, mainly lines, are needed in the ultraviolet).

Another method that can be used to determine the temperature structure is the analysis of disk-center *absolute intensities* as a function of frequency. With the advent of precise absolute intensity measurements in the ultraviolet from space vehicles, this method has surpassed limb-darkening analyses in importance. An extensive compilation and discussion of existing data is given in (646). The basic procedure employed is to construct models by (1) assuming a temperature distribution, (2) integrating the equation of hydrostatic equilibrium, and (3) solving the coupled transfer and statistical equilibrium equations for the relevant atoms (e.g., H, He, C, Si) to calculate absolute emergent intensities. The assumed $T(h)$ relation is then adjusted to yield an optimum fit to all available data. A sequence of successively more refined models have been constructed in this way, starting with the *Utrecht Reference Photosphere* (283, 239), which was soon displaced by the *Bilderberg Continuum Atmosphere* (BCA) (248). Introduction of ultraviolet space observations led to the highly-successful *Harvard-Smithsonian Reference Atmosphere* (HSRA) (summarized in Table 7-1) (249); subsequent work (645; 646) has produced a new model (“Model M”) that fits a very large body of data extremely well. Improvements in these models are constantly being made, and important advances will result when eclipse data are taken into account and spectrum-line synthesis is pushed to its utmost. It should be recognized that there are still many fundamental difficulties to be faced by these semi-empirical models. One of the most worrisome is that observation reveals that the layers under analysis contain a great deal of small-scale structure and are far from homogeneous. Some difficult questions remain essentially unanswered: To what extend does the value assigned to a physical variable in any layer of a one-dimensional chromospheric model represent the average property of that layer? How large is the fluctuation of any physical property about its mean value? Are fluctuations in various properties correlated? It

TABLE 7-1
*Harvard-Smithsonian Reference Atmosphere**

τ_{5000}	$h(\text{km})$	$T(\text{°K})$	$p_p(\text{dynes cm}^{-2})$	$p_s(\text{dynes cm}^{-2})$	$n_p/[n_p + n(\text{H})]$
1.00 - 8	1860	8930	1.52 - 1	4.82 - 2	5.11 - 1
2.00 - 8	1850	8750	1.54 - 1	4.84 - 2	5.05 - 1
3.16 - 8	1840	8630	1.56 - 1	4.92 - 2	5.06 - 1
6.31 - 8	1830	8450	1.62 - 1	5.10 - 2	5.05 - 1
1.00 - 7	1820	8320	1.69 - 1	5.34 - 2	5.09 - 1
2.00 - 7	1790	8090	1.88 - 1	5.88 - 2	5.02 - 1
3.16 - 7	1760	7910	2.10 - 1	6.44 - 2	4.86 - 1
6.31 - 7	1690	7630	2.76 - 1	7.50 - 2	4.09 - 1
1.00 - 6	1620	7360	3.76 - 1	7.85 - 2	2.89 - 1
2.00 - 6	1430	6720	1.00 + 0	6.85 - 2	7.99 - 2
3.16 - 6	1230	6180	3.79 + 0	6.00 - 2	1.72 - 2
6.31 - 6	947	5590	2.92 + 1	6.80 - 2	2.26 - 3
1.00 - 5	840	5300	6.79 + 1	6.77 - 2	9.13 - 4
2.00 - 5	720	4910	1.87 + 2	4.43 - 2	1.41 - 4
3.16 - 5	654	4660	3.37 + 2	3.96 - 2	2.46 - 5
6.31 - 5	588	4280	6.34 + 2	5.11 - 2	7.65 - 7
1.00 - 4	557	4170	8.68 + 2	6.12 - 2	2.26 - 7
2.00 - 4	515	4205	1.34 + 3	9.09 - 2	2.13 - 7
3.16 - 4	487	4250	1.77 + 3	1.20 - 1	2.47 - 7
6.31 - 4	447	4330	2.65 + 3	1.81 - 1	3.40 - 7
1.00 - 3	420	4380	3.46 + 3	2.37 - 1	4.06 - 7
2.00 - 3	379	4460	5.12 + 3	3.53 - 1	5.43 - 7
3.16 - 3	352	4525	6.65 + 3	4.65 - 1	7.11 - 7
6.31 - 3	311	4600	9.81 + 3	6.86 - 1	8.87 - 7
1.00 - 2	283	4660	1.27 + 4	8.95 - 1	1.09 - 6
2.00 - 2	241	4750	1.87 + 4	1.33 + 0	1.46 - 6
3.16 - 2	212	4840	2.41 + 4	1.78 + 0	2.13 - 6
6.31 - 2	168	5010	3.54 + 4	2.81 + 0	4.42 - 6
1.00 - 1	138	5160	4.56 + 4	3.95 + 0	8.49 - 6
2.00 - 1	92.6	5430	6.61 + 4	7.04 + 0	2.47 - 5
3.16 - 1	63.1	5650	8.31 + 4	1.13 + 1	5.29 - 5
6.31 - 1	22.6	6035	1.12 + 5	2.65 + 1	1.58 - 4
1.00	0.0	6390	1.31 + 5	5.64 + 1	3.65 - 4
2.00	-25.3	7140	1.54 + 5	2.35 + 2	1.55 - 3
3.16	-37.1	7750	1.65 + 5	6.26 + 2	4.04 - 3
6.31	-51.4	8520	1.78 + 5	1.81 + 3	1.11 - 2
10.00	-60.8	8880	1.86 + 5	2.61 + 3	1.67 - 2
20.00	-76.7	9390	2.00 + 6	5.00 + 3	2.79 - 2

* The notation $a \cdot aa - b$ denotes $a \cdot aa \times 10^{-b}$.

Adapted from (249), by permission.

is imperative that we face these queries, for one-dimensional models are the best we can hope to obtain for *stellar chromospheres*, and the validity of such models will be supported or undermined, depending on the answers to the questions posed above for the *solar chromosphere*.

In the corona, which is optically thin, the density structure can be inferred from the brightness of the light scattered by free electrons, after allowing for the effects of integration along the line of sight [see, e.g., (94)]. Coronal density distributions have been tabulated for the equatorial and polar directions, for varying conditions during the solar cycle (477); a typical electron density near the base of the corona is of the order of 10^8 cm^{-3} . That the temperature of the corona must be very high was first recognized by Grotrian, who suggested that two broad, faint absorption features seen in the light scattered by the corona were the very strong Ca II *H*- and *K*-lines of the solar spectrum, "washed out" by electron scattering at very high temperatures. This idea was further supported by the identification, by Elden, of coronal emission lines with transitions in highly ionized atoms—e.g., Fe X ($\lambda 6374$) and [Fe XIV] ($\lambda 5303$). Modern determinations of the temperature in the corona are in agreement with one another and give a typical temperature of about $1.5 \times 10^6 \text{ °K}$, based on (a) thermal line-widths, (b) ionization equilibria (including dielectronic recombination!), and (c) radio-wavelength noise temperatures. X-ray observations reveal intensely hot regions (several millions of degrees) near active regions. For more detailed discussion see (20) and (94).

The temperature distributions shown in Figures 7-31 and 7-32 obviously bear little resemblance to the classical predictions of radiative-convective equilibrium models, and indicate a major breakdown in this approach in the outer layers of the atmosphere. The classical models are adequate to predict the structure of the photospheric layers where the visual continuum and line wings (and very weak lines) are formed; but once the temperature rises outward we must introduce new phenomena and an appropriate heating mechanism (clearly the Cayrel mechanism is not the cause of the rise, for T_e becomes $\gg T_{\text{eff}}$!). It was suggested by Biermann (89) and Schwarzschild (564) that acoustic waves would be generated in the solar convection zone, propagate outward, steepen into shocks, and deposit energy in the material, heating it to high temperatures. A specific mechanism for the production of soundwaves was suggested by Lighthill (394). Subsequent work has also called attention to the role of magnetohydrodynamic waves, for the outer solar atmosphere has a significant magnetic field. In addition, the discovery (393) that large regions on the solar surface oscillate with a period of about 300° provided another source of nonradiative energy that can be tapped to produce heating. The problems of wave generation, propagation, and dissipation are complex and difficult, and a tremendous

amount of work has been done on the subject; an excellent review of the field has been given by Stein and Leibacher (602) [see also (20)]. No fully consistent picture of the heating has yet emerged, but current work indicates that the 5-minute oscillations can indeed heat the upper chromosphere and corona, while shorter periods are required to heat the lower chromosphere. Estimates of the actual amount of mechanical energy input can be made by comparing the semiempirical temperature distributions with radiative-equilibrium models [e.g., (17) or (380)] and calculating the energy required to produce the heating. Difficulties here are that at the temperature minimum (1) T in the empirical models lies *below* T in the radiative models, and (2) small errors in T (say $\pm 100^{\circ}\text{K}$) imply unacceptable errors in the energy input estimates. Much work remains to be done to establish both the empirical and theoretical temperatures to the required accuracy.

The presence of temperature plateaus followed by large jumps in temperature can be understood in terms of thermal stability of the gas. In essence, the gas is heated by mechanical input and cooled by radiative losses, which establish a balance [see (628) and (20)]. The losses are largest just as some atom is becoming nearly completely ionized, and this tends to provide a thermostatic action that holds the gas near a definite temperature. Thus, hydrogen provides the dominant cooling in the low chromosphere and keeps the gas at about 7000°K to 8000°K ; after hydrogen becomes too strongly ionized to be effective, losses from He I and He II dominate, and the temperature jumps to about $20,000^{\circ}\text{K}$ to $30,000^{\circ}\text{K}$; and, finally, after helium is strongly ionized, losses from ions of C, N, O, Ne, Mg, and Si dominate, yielding strong cooling rates at temperatures above 10^5°K . In regions of steep gradients, conductive energy transfer also becomes extremely important and the final temperature profile is a result of all these mechanisms operating together.

Given that the sun is a *typical* G-dwarf, it is clear that we must conclude that the *chromosphere-corona-wind phenomenon must be a basic property of stars in general*. There is ample evidence that this is true. Thus, most stars in which convection zones exist show chromospheric emission features in the Ca II H and K lines (cf. §12-2), with many stars that are younger than the sun showing very active chromospheres. An extreme case is provided by the T-Tauri stars where most of the prominent spectral features arise in a “super-chromosphere” and very dense wind. Mass loss in stellar winds, particularly in the early-type supergiants and WR stars (estimated rates up to 10^{-6} – $10^{-5} M_{\odot}/\text{year}$) is well established. These winds make the solar wind (mass-loss rate $\approx 10^{-14} M_{\odot}/\text{year}$) seem puny by comparison.

The solar atmosphere provides a kind of Rosetta stone that helps us first understand a rich “literature” containing many dramatic stellar phenomena; these phenomena, in turn, extend our knowledge over a broad range of physical conditions. Further, by studying large numbers of stars, we can

hope to infer, from the behavior of the *ensemble*, the time-evolution of the phenomena for an *individual* such as the sun, on otherwise inaccessible time-scales (billions of years). Work at the interface between solar and stellar atmospheres is both active and richly rewarding [see, e.g., (344)], and will unquestionably be synergistic in the development of our concepts of stellar atmospheres.

8

The Line Spectrum: An Overview

Superimposed upon the continuum of a star, we observe discrete spectrum lines, either in absorption or emission. These lines arise from transitions between bound states of atoms and ions in the star's atmosphere. An extremely wide variety of lines is found, from a wide range of atomic and ionic states, leading to very different-looking spectra for different classes of stars. A panoramic view of this variety is best obtained by inspection of actual stellar spectra, particularly as shown in (5) and (465) [see also (261, Chap. 14; 330, Chap. 1)]. One finds that lines in stellar spectra show enormous ranges of strength and striking variations in profile. A close examination shows that the spectra can be arranged into a two-dimensional scheme reflecting primarily the effective temperature and luminosity of the star. It would take us too far afield to describe here the details of this procedure—developed to a high state of refinement by Morgan and his collaborators—or the full implications of the results; therefore the references cited above should be studied carefully. Suffice it to say, the spectrum lines contain a wealth of information concerning the run of physical variables in the star and therefore provide important diagnostic tools for inferring the state of the atmosphere.

Spectral lines are much more opaque in the core than in the wing, and hence can provide a sampling of a wide range of atmospheric depths, from very high layers (seen in the core) to the deepest points observable (the depth of continuum formation). Further, lines are narrow in frequency-width, and hence are sensitive to the effects of velocity fields; they thus provide the means by which motions of the material in stellar atmospheres can be studied.

Moreover, it is clear that the strength of a given line must contain information about the number of atoms absorbing photons along the line of sight, and hence about the abundance of that chemical species in the atmosphere. Thus, by a suitable interpretation, the line spectrum offers the opportunity to perform a quantitative chemical analysis of the material of which stars are composed. This information, in turn, provides valuable clues when we attempt to construct a coherent picture of the structure and evolution of stars, the Galaxy, and the Universe as a whole.

It is, therefore, of prime importance to develop a theoretical framework within which line profiles can be predicted and the desired physical information can be inferred. A great deal of effort has been devoted by many astronomers to this end, and considerable progress has been made. The second part of this book will describe the theoretical techniques that now exist to treat the problem of line-formation. In this chapter a few of the basic aspects of the problem, and a summary of the kinds of information required to solve it will be pointed out to orient our later work.

8-1 Observational Quantities

A line in a stellar spectrum is most completely characterized by its *profile*, which is the observed distribution of energy as a function of frequency. For all stars except the sun, we can observe only the flux integrated over the entire disk of the star. We measure F_v (the flux in the line) relative to F_c (the flux in the continuum) and describe the profile in terms of its *absorption depth*

$$A_v \equiv 1 - (F_v/F_c) \quad (8-1)$$

or *residual flux*

$$R_v \equiv (F_v/F_c) = 1 - A_v \quad (8-2)$$

In the case of the sun, the frequency distribution of the radiation can be observed at each point on the disk. We can then describe the profile in terms of $I_v(0, \mu)$, the emergent specific intensity, in units of the nearby continuum intensity $I_c(0, \mu)$, and write

$$a_v(\mu) \equiv 1 - [I_v(0, \mu)/I_c(0, \mu)] \quad (8-3)$$

or $r_v(\mu) \equiv I_v(0, \mu)/I_c(0, \mu) = 1 - a_v(\mu)$ (8-4)

Information about the *center-to-limb variation* of a profile is extremely valuable because it provides (via the Eddington-Barbier relation) an additional depth-resolution otherwise unavailable, and places important constraints on the theory. Such information is, of course, available *only* for the solar spectrum, and is one of the reasons why the solar spectrum provides an ideal testing-ground for proposed theories of line formation.

Often, because of the low light levels involved, it is not possible to measure a stellar spectrum with sufficient resolution to determine a line-profile in detail, and one must then substitute the integrated line strength—the *equivalent width*—in place of this more detailed information. For stars, where we measure the *flux* from the disk, the equivalent width is defined as

$$W_v \equiv \int_0^\infty A_v dv \quad (8-5)$$

in frequency units (hz), or, more usually, as

$$W_\lambda \equiv \int_0^\infty A_\lambda d\lambda \quad (8-6)$$

in wavelength units (\AA or m\AA). For the *sun*, where angular information is available, we can define, in addition,

$$W_v(\mu) \equiv \int_0^\infty a_v(\mu) dv \quad (8-7)$$

or $W_\lambda(\mu) \equiv \int_0^\infty a_\lambda(\mu) d\lambda \quad (8-8)$

The equivalent width is clearly the width of a perfectly black line with the same area under the continuum level as the line under study (hence the name); obviously W gives a direct measure of the *total energy* in the continuum *removed by the line* (assuming that it is in absorption).

Ideally, one attempts to obtain profiles rather than equivalent widths, for they contain far more information. In particular, it is apparent that there is an infinite number of radically different profiles (each with distinct implications for the structure of the atmosphere) that will produce a given equivalent width. An interpretation based on an equivalent width alone can be misleading (the same remark holds even for profiles!). Nevertheless, approaches exist that use equivalent width information from many lines simultaneously (the *curve of growth*), and these can yield important and reasonably unambiguous results. The actual measurement of the data requires refined instrumental techniques; we shall not discuss these techniques here as they lie beyond the scope of this book, but excellent discussions exist elsewhere [see, e.g., (300, Chaps. 2, 4, and 13)]

8-2 The Physical Ingredients of Line-Formation

As in the case of the continuum, the calculation of the flux in a line requires the solution of a transfer problem, for the observed radiation originates from a wide range of depths, over which the physical properties of the material may vary more or less strongly. Let us inquire here what information is needed to formulate and solve the transfer equation in a spectrum line.

Consider a frequency v , and suppose that we know the continuous absorption and scattering coefficients, κ_v and σ_v , and the line absorption coefficient $\chi_l \phi_v$ as a function of depth. We could then construct the line optical depth scale

$$\tau_v \equiv \int_z^{z_{\max}} (\kappa_v + \sigma_v + \chi_l \phi_v) dz' \quad (8-9)$$

and continuum optical depth scale

$$\tau_c \equiv \int_z^{z_{\max}} (\kappa_c + \sigma_c) dz' \quad (8-10)$$

Then if, in addition, we knew the run of the source function S_v , we could immediately calculate

$$F_v = 2 \int_0^\infty S_v(\tau_v) E_2(\tau_v) d\tau_v \quad (8-11)$$

$$F_c = 2 \int_0^\infty S_c(\tau_c) E_2(\tau_c) d\tau_c \quad (8-12)$$

and hence A_v from equation (8-1), and W_λ from equation (8-6) if desired. Of course, in practice, it is the *source function that must be determined*. Only in the trivial case of LTE is S_v known beforehand (namely $S_v = B_v$). As was described in Chapters 2 and 7, in the general (non-LTE) situation, the source function and optical depth depend explicitly upon the occupation numbers of the particular levels involved, but these, in turn, depend upon the radiation field, and hence, ultimately, upon the source function. Thus what is required, as was the case for the continuum, is a *simultaneous, self-consistent* solution of the coupled transfer and statistical equilibrium equations.

Before we attempt to do this, however, some insight into the kinds of information that will be required to attack the problem can be gained from the following phenomenological arguments. Consider the propagation of photons in a line with overlapping continuous absorption and scattering. Some photons will interact with the continuum, others with the line opacity. Some of the photons absorbed in the line will be scattered, and in general will suffer redistribution in frequency and angle according to the redistribution function $R(v', n'; v, n)$. Others may be destroyed by collisional de-excitations or transitions to other levels. Photons may be introduced into the line by collisional excitations or by transitions into the upper level from other levels, with subsequent cascade to the lower level. A transfer equation

accounting for these processes will be of the form

$$\begin{aligned}\mu(\partial I_v / \partial z) = & -(\kappa_v + \sigma_v + \chi_i \phi_v) I_v + \kappa_v B_v + \sigma_v J_v \\ & + \tilde{\gamma} \chi_i \oint (d\omega / 4\pi) \int_0^\infty dv' I(v', n') R(v', n'; v, n) + \tilde{\epsilon} \chi_i \phi_v B_v\end{aligned}\quad (8-13)$$

Here the coupling coefficients $\tilde{\gamma}$ and $\tilde{\epsilon}$ describe, respectively, the fraction of the photons scattered, and those emitted by other processes at a characteristic temperature T_R (not equal, in general, to the local electron temperature).

At this point we can see four important ingredients needed to compute line profiles theoretically.

(a) We must be able to calculate the absorption profile ϕ_v . This will be treated in Chapter 9.

(b) We must be able to specify the coupling between photons and the material via the parameters $\tilde{\gamma}$ and $\tilde{\epsilon}$. In general, the expressions for these quantities may be very complicated and may contain both radiative and collisional rates from the levels giving rise to the line, as well as terms coupling to *other* levels. The specification of $\tilde{\gamma}$ and $\tilde{\epsilon}$ follows from the statistical equilibrium equations for the atom as a whole. We shall examine this aspect of the problem in Chapters 11 and 12.

(c) We must be able to calculate the redistribution function $R(v', n'; v, n)$, and to determine the effects of the details of photon scattering upon the line profile. We shall consider these problems in Chapter 13.

(d) We must be able to solve the resulting transfer problem. Here we will simply refer back to the difference-equation methods discussed in Chapter 6.

In all four of the areas listed above, great strides forward have been made recently. The greatest improvement in our understanding has come in regards to points (b), (c), and (d). In early treatments of line-formation, the question of redistribution was often evaded, and line-scattering was treated as coherent. We know today that this is a poor approximation, and that, in fact, a much better approximation is the opposite extreme assumption of complete redistribution over the line. In the specification of the parameters $\tilde{\gamma}$ and $\tilde{\epsilon}$, the classical approach was sketchy, and led to some serious misconceptions. More modern treatments have brought to light the importance of a clear understanding of these coefficients. In the area of actually solving the transfer equation, important advances have been made possible by application of high-speed, large-capacity computers, using the recently-developed, powerful, numerical techniques.

Finally, it should be noted that in equation (8-13) it has tacitly been assumed that the atmosphere is static. The important effects of velocity fields upon line-formation are discussed in Chapter 14 where several different techniques for solving the transfer equation in moving media are described. The role of lines in establishing the *dynamical* state of the atmosphere is discussed in Chapter 15 in the sections on radiatively-driven stellar winds.

The Line Absorption Profile

The profiles of lines in stellar spectra contain information about both the physical conditions and the abundances of chemical elements in the stellar atmosphere. Therefore they provide extremely valuable diagnostic tools and must be exploited as fully as possible. To carry out an analysis of observed line profiles one needs to know how the distribution of opacity with frequency in the line—the line *absorption profile*—depends upon local conditions of density, temperature, etc. For an isolated atom with levels having essentially infinite lifetimes, the spectral lines would be almost perfectly sharp; but in reality there are several different mechanisms that produce an indefiniteness in the energy levels of real atoms in a plasma, and thereby lead to *line broadening*.

The first line-broadening mechanism to be considered below is *natural* (or *radiation*) *damping*, which refers to the line width produced by the finite lifetime of the atomic levels set by their decay via the radiation process itself. Natural damping occurs even for a solitary, isolated atom. If the atom is imbedded in a plasma, then there will be an additional *pressure broadening* of the line caused by perturbations of the radiated wavetrain through collisions with other atoms, or charged particles, in the gas.

Classically, pressure broadening is described in terms of two limiting approximate theories. The first of these is known as *impact theory*. Here the radiating atom is considered to be an oscillator that suffers a collision that occurs essentially instantaneously, and that interrupts the radiation wavetrain with a sudden phase shift, or by inducing a transition. These collisions thus cause the radiator to "start" and "stop" in intervals of finite duration, leading thereby to a frequency spread in the radiated wavetrain, and to a shift of the line away from its unperturbed frequency. The alternative approach is the *statistical theory* in which we consider the atom to be radiating in a field produced by an ensemble of perturbers. This field will fluctuate statistically about some mean value as a result of motions of the perturbers. At a given value of the field, the energy levels of the radiating atom are shifted slightly, and, correspondingly, the frequency of the line is altered. The intensity of the radiation at any specified frequency shift is taken to be proportional to the probability with which a perturbation of the appropriate field strength occurs.

The basic limits on the *classical* theories of pressure broadening are set by their inability to account for the actual *structure* of the radiating atom, or for *transitions* produced by the collisions with the perturbing particles. Both of these defects are overcome in the *quantum theory* of pressure broadening, which yields results in good agreement with experimental determinations.

Finally, we must account for the fact that in a stellar atmosphere we observe an ensemble of atoms moving with a velocity distribution along the line of sight. The profile for each atom is Doppler shifted according to its line-of-sight velocity, and the profile seen from the entire ensemble is a superposition of these shifted atomic profiles, calculated by a convolution with the velocity distribution.

The theory of line broadening has progressed enormously in the past decade, and reliable calculations now exist for many profiles of fundamental astrophysical interest. The quantum theory has become very refined, but also quite complicated. As excellent treatises exist on the general subject of line broadening [see particularly (264; 268; 629)], only a brief summary of the most important results will be presented in this chapter; for further details one should consult the books just cited or the research literature, for which references are compiled in (228; 229; 232).

9-1 The Natural Damping Profile

ENERGY SPECTRA, POWER SPECTRA, AND THE AUTOCORRELATION FUNCTION

We first derive some basic relations that will be needed in later discussions. Consider a time-varying oscillation of amplitude $f(t)$. The *Fourier*

transform $F(\omega)$ is defined to be

$$F(\omega) \equiv \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt \quad (9-1)$$

which satisfies the fundamental *reciprocity relation*

$$f(t) = (2\pi)^{-1} \int_{-\infty}^{\infty} F(\omega) e^{i\omega t} d\omega \quad (9-2)$$

The quantity

$$E(\omega) \equiv (2\pi)^{-1} F^*(\omega) F(\omega) \quad (9-3)$$

is called the *energy spectrum* of the oscillator. This designation derives from the fact that

$$\int_{-\infty}^{\infty} E(\omega) d\omega = (2\pi)^{-1} \int_{-\infty}^{\infty} F^*(\omega) F(\omega) d\omega = \int_{-\infty}^{\infty} f^*(t) f(t) dt \quad (9-4)$$

which may be verified by direct calculation, using equation (9-1). If $f(t)$ were the voltage across a one-ohm resistor, then $f^*(t)f(t)$ would be the instantaneous power delivered to the resistor, and the integral over all time gives the total energy. Thus, $E(\omega)$ is a direct measure of the energy in the wavetrain at frequency ω .

In many cases, the energy spectrum itself is not used, but rather the energy delivered per unit time—i.e., the *power spectrum* $I(\omega)$, which is defined as

$$I(\omega) = \lim_{T \rightarrow \infty} (2\pi T)^{-1} \left| \int_{-T/2}^{T/2} f(t) e^{-i\omega t} dt \right|^2 \quad (9-5)$$

However, for oscillations of finite duration (or with, say, an exponential decay), the power spectrum will be zero because on the average over an *infinite* time interval, the *finite* total energy emitted yields zero power. In these cases, which are of practical interest, we use the energy spectrum itself, assuming that we observe an *ensemble* of oscillators *created at a constant rate* with *random phases*; a finite power then results, with a frequency distribution proportional to the energy spectrum of an individual oscillator.

In certain situations it is not possible to calculate the power spectrum directly, using equation (9-5). It is then valuable to make use of the *autocorrelation function*

$$\Phi(s) \equiv \lim_{T \rightarrow \infty} T^{-1} \int_{-T/2}^{T/2} f^*(t) f(t + s) dt \quad (9-6)$$

from which the power spectrum is obtained by the relation

$$I(\omega) = (2\pi)^{-1} \int_{-\infty}^{\infty} \Phi(s) e^{-i\omega s} ds \quad (9-7)$$

as may be verified by direct calculation, using a limiting procedure in the integration over s . The autocorrelation function provides a powerful tool for calculating power spectra of a radiating atom perturbed by collisions.

THE DAMPED CLASSICAL OSCILLATOR

The simplest picture one can construct of the process of emission in a line is to consider the atom to be a classical oscillator. In §4-2 it was shown that the profile from a *driven* damped oscillator is Lorentzian, by a calculation of the power emitted by the oscillator. We shall apply the techniques outlined above to a decaying oscillator to show that the same profile is obtained. From classical electromagnetic theory the equation of motion [cf. equation (4-27)] for a radiating oscillator is

$$\ddot{x} = -\omega_0^2 x - \gamma \dot{x} \quad (9-8)$$

where γ is the *classical damping constant*

$$\gamma \equiv (2e^2\omega_0^2/3mc^3) \quad (9-9)$$

The radiation damping term is numerically quite small, and may be estimated using the undamped solution $x = x_0 \exp(i\omega_0 t)$, which yields

$$\ddot{x} = -(\omega_0^2 + i\gamma\omega_0)x \quad (9-10)$$

Neglecting terms in γ^2 , the solution of equation (9-10) is

$$x = x_0 e^{i\omega_0 t} e^{-\gamma t/2} \quad (9-11)$$

which is an exponentially damped oscillation. Calculating the Fourier transform, assuming the oscillation starts abruptly at $t = 0$, we find

$$F(\omega) = x_0 \int_0^\infty e^{-i(\omega-\omega_0)t} e^{-\gamma t/2} dt = x_0 / [i(\omega - \omega_0) + \frac{1}{2}\gamma] \quad (9-12)$$

The energy spectrum of the oscillator is then

$$E(\omega) = (x_0^2/2\pi) [(\omega - \omega_0)^2 + (\frac{1}{2}\gamma)^2]^{-1} \quad (9-13)$$

The power spectrum of an ensemble of such oscillators created continuously with random phases is proportional to $E(\omega)$, hence the profile, normalized such that

$$\int_{-\infty}^{\infty} I(\omega) d\omega = 1 \quad (9-14)$$

is

$$I(\omega) = (\gamma/2\pi) [(\omega - \omega_0)^2 + (\frac{1}{2}\gamma)^2]^{-1} \quad (9-15)$$

The damped classical oscillator thus yields a *Lorentz profile* with a full half-intensity width γ . In wavelength units this width is

$$\Delta\lambda_c = (2\pi c\gamma/\omega^2) = (4\pi e^2/3mc^2) = 1.2 \times 10^{-4} \text{ \AA} \quad (9-16)$$

This width is much smaller than those observed in laboratory or stellar spectra. We must therefore develop a more general picture of the radiation process.

QUANTUM MECHANICAL CALCULATION

A quantum mechanical analogue of the damped oscillator is obtained by assuming the radiation arises in transitions by an atom from an excited state of finite lifetime to the ground state. Following Wigner and Weisskopf, we write the probability of finding an atom in the excited state j as

$$P_j(t) = \psi_j^* \psi_j e^{-\Gamma t} \quad (9-17)$$

where $\Gamma = A_{ji}$, the spontaneous emission rate. Then the time development of the wave function of the state is

$$\psi_j(r, t) e^{-\Gamma t/2} = u_j(r) e^{-iE_j t/\hbar} e^{-\Gamma t/2} = u_j(r) e^{-(i\omega_j + \frac{1}{2}\Gamma)t} \quad (9-18)$$

Consistent with the *uncertainty principle*, we consider that the decaying state j (with characteristic lifetime Δt_j) no longer has a perfectly defined energy E_j , but is, rather, a superposition of states with energies spread about E_j (with a characteristic width $\Delta E_j \sim \hbar/\Delta t_j$). From the fundamental reciprocity relations of quantum mechanics, the *amplitude* of the energy distribution is given by the Fourier transform of the time dependence of the wave function, and the *probability distribution* of energy states is given by the square of the amplitude. Thus calculating the Fourier transform of equation (9-18), it is clear that the result must be of the same form as that derived from equation (9-11), and by analogous arguments we obtain finally

$$I(\omega) = (\Gamma/2\pi) [(\omega - \omega_0)^2 + (\frac{1}{2}\Gamma)^2]^{-1} \quad (9-19)$$

From equation (9-17) we see that Γ is to be interpreted as the reciprocal of the mean lifetime of the upper state. If *several* transitions out of the upper state U are possible, then

$$\Gamma_U = \sum_{i < U} A_{Ui} \quad (9-20)$$

is the appropriate width of the state. Suppose the line under consideration arises from a transition between two *excited* states, so that the lower state L also has a width Γ_L given by an equation analogous to (9-20). The line profile must, in general, reflect the width of *both* states. We assume that the distribution of substates of each level around its nominal energy is given by a Lorentz profile with the appropriate Γ . Let $\delta \equiv \Gamma/2$ for either level, and let $x = (E - E_0)/\hbar$ be the frequency displacement of a particular substate from the nominal energy E_0 . We assume that the probability of ending in a particular substate x' of the lower level is independent of the substate x of the upper level from which the transition starts; then the joint probability of starting in substate x and ending in substate x' is

$$p(x, x') = (\delta_L \delta_U / \pi^2) [(x^2 + \delta_U^2)(x'^2 + \delta_L^2)]^{-1} \quad (9-21)$$

If we restrict attention to transitions producing radiation of a specific frequency ω , then x and x' must be related by $\omega_0 + x - x' = \omega$, or writing $x_0 \equiv \omega - \omega_0$,

$$x' = x - x_0 \quad (9-22)$$

The total intensity at ω is obtained by summing over all upper substates x , with x' fixed by equation (9-22), that is

$$\begin{aligned} I(\omega) &= \int_{-\infty}^{\infty} p(x, x - x_0) dx \\ &= \frac{\delta v \delta_L}{\pi^2} \int_{-\infty}^{\infty} \frac{dx}{(x^2 + \delta_v^2)[(x - x_0)^2 + \delta_L^2]} \end{aligned} \quad (9-23)$$

This integral may be evaluated by contour integration using the *residue theorem*, taking into account the poles at $z = \pm i\delta_v$ and $z = x_0 \pm i\delta_L$. Performing the integration one again finds $I(\omega)$ is given by equation (9-19) but now with the damping width given by

$$\Gamma \equiv \Gamma_L + \Gamma_v \quad (9-24)$$

Hence the profile is *Lorentzian with a half-intensity width equal to the sum of the half-intensity widths of both levels*.

Exercise 9-1: Evaluate equation (9-23) by contour integration and verify equation (9-24).

The Lorentz profiles calculated above are, strictly speaking, *emission profiles*. If, however, we assume *detailed balancing*, then the *absorption profile* will have the *same form*. To convert to absorption cross-sections per atom we recall from equations (4-34) and (4-35) that

$$\int_{-\infty}^{\infty} \alpha_v dv = (\pi e^2 / mc) f \quad (9-25)$$

Thus using a profile of the form of equation (9-19) and converting to ordinary frequency units, the absorption cross-section is

$$\alpha_v = \left(\frac{\pi e^2}{mc} \right) f \frac{(\Gamma/4\pi^2)}{(v - v_0)^2 + (\Gamma/4\pi)^2} \quad (9-26)$$

Radiation damping is of primary importance for strong lines in low-density media, for example $L\alpha$ in interstellar space. In most cases of interest in stellar atmospheres, however, the line is formed in regions where the density of perturbing atoms, ions, and electrons is high enough that pressure broadening is significant (or dominant).

9-2 Effects of Doppler Broadening: The Voigt Function

When one observes a line in a stellar atmosphere (or a laboratory plasma), one sees the combined effects of absorption by all atoms in the ensemble. Each atom will have a velocity along the line of sight, measured in the observer's frame, and the intrinsic profile of that atom will be Doppler-shifted a corresponding amount in frequency. If the damping process producing the intrinsic profile of each atom is *uncorrelated* with its velocity, then the shifted profiles may be superimposed to yield the total absorption cross-section by a simple folding procedure.

Assuming the plasma is characterized by a kinetic temperature T , the velocity distribution is Maxwellian, so that the probability of finding an atom with a line-of-sight velocity ξ on the range $(\xi, \xi + d\xi)$ is

$$W(\xi) d\xi = (\pi^{1/2} \xi_0)^{-1} \exp(-\xi^2/\xi_0^2) d\xi \quad (9-27)$$

where $\xi_0 = (2kT/m)^{1/2} = 12.85 (T/10^4 A)^{1/2}$ km s⁻¹, where A is the atom's atomic weight. Then, if we observe at frequency v , an atom with velocity component ξ is absorbing at frequency $v[1 - (\xi/c)]$ in its own frame, and the absorption coefficient for that atom is $\alpha(v - \xi v/c)$. The total absorption coefficient at frequency v is thus given by the convolution integral

$$\alpha_v = \int_{-\infty}^{\infty} \alpha(v - \xi v/c) W(\xi) d\xi \quad (9-28)$$

Equation (9-28) can be applied to any absorption profile to allow for the effects of *Doppler broadening*; for the remainder of this section consideration will be restricted to the case where the intrinsic profile is Lorentzian.

Substituting equations (9-26) and (9-27) into equation (9-28), and defining

$$v \equiv (v - v_0)/\Delta v_D \quad (9-29)$$

$$y \equiv (\Delta v/\Delta v_D) = (\xi/\xi_0) \quad (9-30)$$

$$\text{and } a \equiv (\Gamma/4\pi \Delta v_D) \quad (9-31)$$

where Δv_D is the *Doppler width* of the line

$$\Delta v_D \equiv (\xi_0 v_0/c) \quad (9-32)$$

we find the absorption coefficient can be written as

$$\alpha_v = (\pi^{1/2} e^2 f / mc \Delta v_D) H(a, v) \quad (9-33)$$

$$\text{where } H(a, v) \equiv \frac{a}{\pi} \int_{-\infty}^{\infty} \frac{e^{-y^2} dy}{(v - y)^2 + a^2} \quad (9-34)$$

is known as the *Voigt function*. In deriving equation (9-34), the approximation that $\xi v/c \approx \xi v_0/c$ (appropriate in a stellar atmosphere) was made. Extensive tables of $H(a, v)$ are given in (219) and (314).

General methods of computing the Voigt function are described in (314; 527; 528). The usual case of astrophysical interest is when $a \ll 1$; in this limit one can develop a convenient expression for $H(a, v)$ as a power series in the parameter a as follows. Using the Laplace transform of the cosine function

$$\int_0^\infty e^{-ax} \cos bx dx = a/(a^2 + b^2) \quad (9-35)$$

and the addition rule for cosines, the Voigt function can be written as

$$H(a, v) = \pi^{-1} \int_0^\infty dx e^{-ax} \cos vx \int_{-\infty}^\infty dy e^{-y^2} \cos xy \quad (9-36)$$

But the cosine transformation of the Gaussian is

$$\int_{-\infty}^\infty e^{-y^2} \cos xy dy = \sqrt{\pi} e^{-x^2/4} \quad (9-37)$$

from which we see that

$$H(a, v) = \pi^{-\frac{1}{2}} \int_0^\infty e^{-(ax+(x^2/4))} \cos vx dx \quad (9-38)$$

Assuming that $a \ll 1$, we can replace e^{-ax} by its power series, and integrating term by term we find

$$H(a, v) = \sum_{n=0}^{\infty} a^n H_n(v) \quad (9-39)$$

$$\text{where } H_n(v) \equiv [(-1)^n / \pi^{\frac{n}{2}} n!] \int_0^\infty e^{-x^2/4} x^n \cos vx dx \quad (9-40)$$

From equation (9-37) it follows immediately that

$$H_0(v) = e^{-v^2} \quad (9-41)$$

Integrating by parts, the first odd-order term can be written as

$$H_1(v) = (-2/\pi^{\frac{1}{2}}) \left(1 - v \int_0^\infty e^{-x^2/4} \sin vx dx \right) \quad (9-42)$$

and using the sine transform of the Gaussian

$$\int_0^\infty e^{-y^2} \sin 2vy dy = e^{-v^2} \int_0^v e^{y^2} dy \equiv F(v) \quad (9-43)$$

we have

$$H_1(v) = (2/\pi^{\frac{1}{2}})[2vF(v) - 1] \quad (9-44)$$

The function $F(v)$ is known as *Dawson's integral*; efficient techniques for computing $F(v)$ are given in (170). The functions $H_n(v)$ are tabulated in (281) and (11, 325) for $n \leq 4$.

Exercise 9-2: (a) Show that the Voigt function has the normalization

$$\int_{-\infty}^\infty H(a, v) dv = \pi^{\frac{1}{2}}$$

(b) Prove the relation $H_n(v) = -[a^2 H_{n-2}(v)/dv^2]/[n(n-1)]$ from which higher-order terms can be generated by recursion from $H_0(v)$ and $H_1(v)$. (c) Write explicit formulae for $H_2(v)$, $H_3(v)$, and $H_4(v)$, expressing $H_3(v)$ in terms of $F(v)$. (d) Show that for $(a^2 + v^2) \gg 1$, $H(a, v) \approx (\pi^{-\frac{1}{2}} a)/(a^2 + v^2)$.

In view of the result of Exercise 9-2(d), we see that for $v \gg 1$, $H(a, v) \approx a/(\pi^{\frac{1}{2}} v^2)$. Thus a schematic representation of the Voigt function is

$$H(a, v) \sim e^{-v^2} + a/(\pi^{\frac{1}{2}} v^2) \quad (9-45)$$

where the first term applies in the line core, $v \leq v^*$, and the second in the line wing, $v \geq v^*$, the quantity v^* being chosen such that the two terms are equal. The line core is clearly dominated by Doppler broadening, while the line wings are dominated by the damping profile.

9-3 Collision Broadening: Classical Impact Theory

THE WEISSKOPF APPROXIMATION

The simplest classical impact theory has its origins in an analysis by Lorentz, who considered the atom to be a radiating oscillator that suffers changes in phase during encounters with perturbing particles. It is assumed that the collisions occur between the radiating atom and a single perturber, one at a time. The collisions are assumed to occur essentially instantaneously, so that the wavetrain suffers an instantaneous phase dislocation that, in effect, terminates it. During the time between collisions the atom is assumed to be unperturbed. Thus suppose that the time between two successive collisions is T , and that in this interval the radiator emits a monochromatic wavetrain $f(t) = \exp(i\omega_0 t)$. The Fourier transform of this finite wavetrain is

$$F(\omega, T) = \int_0^T e^{i(\omega_0 - \omega)t} dt = \frac{\exp[i(\omega - \omega_0)T] - 1}{i(\omega - \omega_0)} \quad (9-46)$$

The energy spectrum $E(\omega, T)$ of this wavetrain is given by substitution of $F(\omega, T)$ into equation (9-3).

In general there is not a unique time interval between collisions; rather, these intervals are distributed probabilistically about some mean value. If

the collisions occur as a result of a random-walk process, and the *mean time between collisions* is τ , then the probability that the interval between two successive collisions lies on the range $(T, T + dT)$ is

$$W(T) dT = e^{-T/\tau} (dT/\tau) \quad (9-47)$$

Hence averaging over all collision times T , we obtain a mean energy spectrum

$$E(\omega) \equiv \langle E(\omega, T) \rangle_T = (2\pi)^{-1} \int_0^\infty F^*(\omega, T) F(\omega, T) W(T) dT \quad (9-48)$$

Computation of the integral, with normalization, yields

$$E(\omega) = \frac{(1/\pi\tau)}{(\omega - \omega_0)^2 + (1/\tau)^2} = \frac{(\Gamma/2\pi)}{(\omega - \omega_0)^2 + (\Gamma/2)^2} \quad (9-49)$$

Exercise 9-3: Derive equation (9-49).

The collision broadening theory described above again yields a Lorentz profile (a result of assuming that all the collisions are distinct), with a damping parameter $\Gamma = 2/\tau$. To complete the theory we must obtain an estimate of τ . As was done for the radiation-damped oscillator, we take the profile of an ensemble of randomly phased oscillators, continuously created, to be proportional to the energy spectrum of a single oscillator [averaged over all times, as given in equation (9-49)]. If both radiation and collision damping occur, with widths Γ_R and Γ_C respectively, and are assumed to be completely uncorrelated, then the profile is a convolution of the two Lorentz profiles. By an analysis similar to that leading to equation (9-24), one may readily show that the combined profile is again Lorentzian with a total width $\Gamma = \Gamma_R + \Gamma_C$. The effects of Doppler broadening can be taken into account as in §9-2, by using a Voigt profile with the appropriate total damping width.

We must now calculate the mean collision time τ . If the radiating atoms and perturbing particles have atomic weights A_r and A_p , respectively, and both have a Maxwellian velocity distribution at a temperature T , then their average relative velocity is

$$v = \langle v^2 \rangle^{1/2} = [(8kT/\pi m_H)(A_r^{-1} + A_p^{-1})]^{1/2} \quad (9-50)$$

Assuming that the effective impact parameter of the collisions responsible for the broadening is ρ_0 , we then have

$$\tau^{-1} = \pi \rho_0^2 N v \quad (9-51)$$

and

$$\Gamma = 2\pi \rho_0^2 N v \quad (9-52)$$

where N is the perturber density. We must now determine ρ_0 .

Following Weisskopf (661) we assume that (a) the perturber is a *classical particle*; (b) the perturber moves with constant velocity past the atom on a straight-line path with *impact parameter* ρ ; (c) the *interaction* between atom and perturber is described approximately by

$$\Delta\omega = C_p/r^p \quad (9-53)$$

where $r(t) = (\rho^2 + v^2 t^2)^{1/2}$, $t = 0$ occurring at the point of closest approach; and (d) *no transitions* in the atom are produced by the action of the perturber. The validity of these assumptions will be considered later. The form of the interaction in equation (9-53) is only approximate but holds over a fairly wide range of distances. The value of the exponent p depends upon the nature of the interaction. Values of astrophysical interest and the interaction they represent are as follows: $p = 2$, linear Stark effect (hydrogen + charged particle); $p = 3$, resonance broadening (atom A + atom A); $p = 4$, quadratic Stark effect (nonhydrogenic atom + charged particle); $p = 6$, van der Waals force (atom A + atom B). The interaction constant C_p must be calculated from quantum theory or measured by experiment.

The *phase shift* induced by the perturbation is

$$\eta(t) = \int_{-\infty}^t \Delta\omega(t') dt' = C_p \int_{-\infty}^t (\rho^2 + v^2 t'^2)^{-p/2} dt' \quad (9-54)$$

The *total phase shift* $\eta(\rho) \equiv \eta(t = \infty)$ is found directly to be

$$\eta(\rho) = C_p \psi_p / v \rho^{p-1} \quad (9-55)$$

$$\text{where } \psi_p = \pi^{1/2} \Gamma[\frac{1}{2}(p-1)] / \Gamma(\frac{1}{2}p) \quad (9-56)$$

Here Γ denotes the usual gamma function; for $p = (2, 3, 4, 6)$ one finds $\psi_p = (\pi, 2, \pi/2, 3\pi/8)$.

We now assume that only those collisions that produce a total phase shift greater than some critical value η_0 are effective in broadening the line. The effective impact parameter for such collisions is thus

$$\rho_0 = (C_p \psi_p / \eta_0 v)^{1/(p-1)} \quad (9-57)$$

and the corresponding value for the damping constant is

$$\Gamma = 2\pi N v (C_p \psi_p / \eta_0 v)^{2/(p-1)} \quad (9-58)$$

Weisskopf arbitrarily adopted $\eta_0 = 1$ as the critical phase shift; with this choice we obtain the *Weisskopf radius* ρ_w from equation (9-57) and the Weisskopf damping parameter Γ_w from equation (9-58).

If C_p is given, the theory described above yields a definite value for Γ , and the results are found to be of the right order of magnitude. Yet there remain serious defects in it. (a) The choice $\eta_0 = 1$ is arbitrary, and there is no means of determining a priori the correct value of η_0 to be used. (b) The theory does not account for the collisions that produce small phase shifts even though the number of such collisions increases as ρ^2 . (c) The theory fails to predict the existence of a *line shift*; as will be shown below this failure arises from the omission of weak collisions, as mentioned in (b).

THE LINDHOLM APPROXIMATION

A significant improvement in the classical impact theory was made by Lindholm (397; 398) and Foley (221). In this approach we consider the radiator to have an *instantaneous frequency* $\omega(t)$ which, because of perturbations, differs from the *nominal frequency* ω_0 by an amount $\Delta\omega(t)$. Then we write

$$f(t) = \exp \left[i\omega_0 t + i \int_{-\infty}^t \Delta\omega(t') dt' \right] \equiv e^{i[\omega_0 t + \eta(t)]} \quad (9-59)$$

where $\eta(t)$ is the *instantaneous phase* of the oscillator. To obtain the line profile, we calculate the autocorrelation function $\Phi(s)$ defined by equation (9-6). Let $\phi(s)$ be defined by $\phi(s) \equiv e^{-i\omega_0 s} \Phi(s)$, which eliminates the unperturbed oscillation. Then from equation (9-6),

$$\begin{aligned} \phi(s) &= \lim_{T \rightarrow \infty} T^{-1} \int_{-T/2}^{T/2} e^{-i\omega_0 s} e^{-i[\omega_0 t + \eta(t)]} e^{i[\omega_0(t+s) + \eta(t+s)]} dt \\ &= \lim_{T \rightarrow \infty} T^{-1} \int_{-T/2}^{T/2} e^{i[\eta(t+s) - \eta(t)]} dt \end{aligned} \quad (9-60)$$

Clearly $\phi(s)$ is the time-averaged value of the *additional phase shift* occurring in the time-interval s . For brevity, write

$$\eta(t, s) \equiv \eta(t + s) - \eta(t) \quad (9-61)$$

Then $\phi(s) = \langle \exp[i\eta(t, s)] \rangle_T$ (9-62)

Further, writing $d\phi(s) = \phi(s + ds) - \phi(s)$, we have

$$d\phi(s) = \langle e^{i\eta(t, s)} (e^{i\eta'} - 1) \rangle_T \quad (9-63)$$

where η' denotes the *change in phase* occurring in the time-interval ds as a result of collisions that take place in that interval. The phase *change* cannot be correlated with the *current value* of the phase if the collisions occur at random. Thus the average of the product can be replaced by the product of

the averages, i.e.,

$$d\phi(s) = \langle e^{i\eta(t, s)} \rangle_T \langle e^{i\eta'} - 1 \rangle_T = \phi(s) \langle e^{i\eta'} - 1 \rangle_T \quad (9-64)$$

If we can calculate the average of $e^{i\eta'}$, we obtain a differential equation for $\phi(s)$.

By forming the average over a sufficiently long time interval T , the randomly-occurring collisions will happen at all values of ρ with an appropriate statistical frequency. We then invoke the *ergodic hypothesis* to replace the average over time by the appropriate sum over impact parameters. The number of impacts that occur on the range $(\rho, \rho + d\rho)$ in time ds is just $(2\pi\rho d\rho)Nv ds$, hence

$$\langle e^{i\eta'} - 1 \rangle_T \rightarrow \langle e^{i\eta'(\rho)} - 1 \rangle_\rho = 2\pi Nv ds \int_0^\infty [e^{i\eta(\rho)} - 1] \rho d\rho \quad (9-65)$$

The integral in equation (9-65) has both a real and imaginary part, so we write

$$\langle e^{i\eta'(\rho)} - 1 \rangle_\rho = -Nv ds (\sigma_R - i\sigma_I) \quad (9-66)$$

$$\text{where } \sigma_R \equiv 2\pi \int_0^\infty [1 - \cos \eta(\rho)] \rho d\rho = 4\pi \int_0^\infty \sin^2 [\frac{1}{2}\eta(\rho)] \rho d\rho \quad (9-67)$$

$$\text{and } \sigma_I \equiv 2\pi \int_0^\infty \sin \eta(\rho) \rho d\rho \quad (9-68)$$

Combining equations (9-64), (9-65), and (9-66) and solving the resulting differential equation with the initial condition $\phi(0) = 1$ we find

$$\phi(s) = \exp[-Nv(\sigma_R|s| - i\sigma_I s)] \quad (9-69)$$

Finally, calculating the intensity from equation (9-7) and normalizing the profile we obtain

$$I(\omega) = \frac{(Nv\sigma_R/\pi)}{(\omega - \omega_0 - Nv\sigma_I)^2 + (Nv\sigma_R)^2} \quad (9-70)$$

Exercise 9-4: Verify equations (9-69) and (9-70).

Thus Lindholm theory yields a Lorentz profile with a damping width

$$\Gamma = 2Nv\sigma_R \quad (9-71)$$

and a *line shift*

$$\Delta\omega_0 = Nv\sigma_I \quad (9-72)$$

The prediction of a shift is in agreement with experiment, where such shifts are observed. Quantum theory yields a profile of the same form as equation (9-70), and gives explicit expressions for Γ and $\Delta\omega_0$ in terms of matrix elements of the perturbing potential and transitions within the atom. As we

shall see below, Lindholm theory yields a unique value of $\Gamma/\Delta\omega_0$ for each choice of p ; quantum theory shows that this ratio actually varies over a moderate range as T and N vary, and is different for each line. The effects of Doppler broadening are taken into account by using a Voigt profile with the appropriate damping parameter and shifted from its rest frequency by an amount $\Delta\omega_0$.

The dominant contributions to σ_R and σ_I come from quite different ranges of impact parameter ρ . From equation (9-55) we note that $\eta(\rho) \propto \rho^{-(p-1)}$. Thus for $(\rho/\rho_w) > 1$, the integrand of σ_R rapidly drops to zero [see (629, 16) or (638, 305)], and the dominant contribution to the line broadening comes from (strong) collisions inside the Weisskopf radius—i.e., $(\rho/\rho_w) < 1$. In contrast, for σ_I , the integrand for $(\rho/\rho_w) < 1$ fluctuates in sign, and averages to nearly zero. Thus the dominant contribution to the line shift comes from (weak) collisions outside the Weisskopf radius. It is easy to understand physically how the shift arises. The very weak collisions ($\eta \ll 1$, $\rho \gg \rho_w$) are extremely numerous and occur at an essentially constant rate, yielding an average phase change per unit time of

$$\bar{\eta} = 2\pi N v \int_{\rho^*}^{\infty} \eta(\rho) \rho d\rho \quad (9-73)$$

where ρ^* is chosen to assure that $\eta(\rho^*) \ll 1$. But as can be seen from equation (9-59), the rate of change of phase is by definition a change $\Delta\omega$ in the oscillator's frequency.

SPECIFIC CASES

Lindholm theory has been most widely applied in astrophysical work for the cases $p = 3, 4$, and 6 . For these cases the integrals in equations (9-67) and (9-68) can be evaluated explicitly to yield the values listed in Table 9-1 [see (629, 14) for details]. The last line of the table gives the value of η_0 that, when inserted into the Weisskopf formula [equation (9-58)], gives the Lindholm Γ . As η_0 is always less than unity, it can be seen that the Weisskopf formula always leads to too small a value of Γ .

TABLE 9-1
Results of Lindholm Theory

p	3	4	6
Γ	$2\pi^2 C_3 N$	$11.37 \left\{ \begin{array}{l} \\ 9.85 \end{array} \right\} C_4^{-\frac{1}{2}} v^{\frac{1}{2}} N$	$8.08 \left\{ \begin{array}{l} \\ 2.94 \end{array} \right\} C_6^{-\frac{1}{2}} v^{\frac{1}{2}} N$
$\Delta\omega_0$			
$\Gamma/\Delta\omega_0$		1.16	2.75
η_0	0.64	0.64	0.61

Resonance broadening, $p = 3$, is of importance mainly for collisions of hydrogen atoms with one another. As the atmosphere must be hot enough that hydrogen is excited to the $n = 2$ state (to produce the observable Balmer lines) but cool enough that it is not dominantly ionized, resonance broadening effects are of interest for solar-type stars. The interaction constant C_3 in equation (9-53) for level n is

$$C_3 = (e^2 f_{1n}/2m\omega_{1n}) \quad (9-74)$$

[see (662; 112, 231)]. A quantum mechanical calculation gives a Γ slightly different from the Lindholm value, namely

$$\Gamma_3 = (16\pi n_1 C_3/3) = (4n_1 e^2 f_{1n}/3mv_{1n}) \quad (9-75)$$

For the hydrogen lines there is no shift $\Delta\omega_0$ because individual Stark components split symmetrically about line center (see §9-4) and the shift is identically zero. Resonance broadening is most significant for the lowest members of a series where Stark broadening is the smallest. The effects of resonance broadening have been shown to be important in the solar $H\alpha$ line but negligible for higher series members (146).

Quadratic Stark effect, $p = 4$, is important for the broadening of lines of nonhydrogenic atoms and ions by impacts with charged particles (electrons), and is the dominant pressure-broadening mechanism for these lines in the atmospheres of early-type stars. In applications of the classical Lindholm theory the interaction constant C_4 was typically estimated from experimental measures of line shifts in electric fields or from time-independent perturbation theory for the quadratic Stark effect [see (11, 319–320) or (638, 326–328) for examples of this procedure]. The resulting damping widths are usually much too small, however, because the Lindholm approximation assumes the collisions are adiabatic (i.e., do not cause transitions in the radiating atom); this assumption is frequently poor, and accurate quantum mechanical calculations including nonadiabatic effects (cf. §9-5) yield much larger line-widths.

Van der Waals interactions ($p = 6$) of non-hydrogenic atoms in collisions with neutral hydrogen atoms is the major source of pressure broadening in solar-type stars. The usual classical treatment accounts for the dipole–dipole interaction term in the potential, and yields [cf. (629, 91–97; 638, 331–334)]

$$\Delta\omega = C_6/r^6 = e^2 \alpha a_0^2 [R_u^{-2} - R_i^{-2}] / (\hbar r^6) \quad (9-76)$$

where α is the polarizability of hydrogen and R^2 is the mean square radius of the two levels. Quantum mechanical results are sometimes available for R^2 ; if not, hydrogenic estimates are used. Using C_6 determined in this manner, Γ can be computed from Lindholm theory. When this is done (e.g., for lines of Fe I), it is found that the predicted values are much too small, by factors of 5 to 30 (382). Quantum mechanical calculations that again

employ only the dipole-dipole term do not lead to large increases in Γ [e.g., (264; 98; 86; 599)], which points to the breakdown of the dipole-dipole approximation rather than other theoretical problems [see also (301; 302; 541)]. Calculations using the more realistic Lennard-Jones potential (303) lead to significantly larger widths. Attempts have been made to include more terms in the expansion of the interaction potential (233); this increases Γ , but still falls short of the observed values. The expansion technique converges slowly, and an alternate approach, applied to the lines of Fe I, evaluated the exact expression for the interaction using scaled hydrogenic wave functions (116) and showed reasonable agreement between predicted and measured Γ 's. Yet another method is based on the proposal that the dominant interaction leading to the line broadening is between the perturber and the valence electron (539; 540). One can then use a Smirnov potential (578) and obtain an expression for the damping parameter. This approach has been used to produce extensive tables (193) giving the parameters α and β in the formula $\Gamma = N\alpha T^\beta$, as functions of the effective quantum numbers n^* of the lower and upper levels, for $s-p$, $p-d$, and $d-f$ transitions.

VALIDITY CRITERIA

(1) An effective impact time τ_s can be defined such that τ_s times the peak value of $\Delta\omega$ for a collision at the effective impact parameter, namely $C_p \rho_0^{-p}$, yields the total phase shift given by equation (9-55). This gives

$$\tau_s = (\psi_p \rho_0 / v) \quad (9-77)$$

For impact theory to be valid, we demand that only one collision at a time occur, so that $\tau_s < \tau = 1/(N\pi\rho_0^2 v)$. Writing $N = 3/(4\pi r_0^3)$ where r_0 is the mean interparticle distance, we find $(\tau_s/\tau) = \frac{1}{2}\psi_p(\rho_0/r_0)^3$. Thus impact theory will be valid only when the density of particles is so low that the Weisskopf radius is small compared to the interparticle distance.

(2) It is clear that, as $\rho \rightarrow \infty$, the effective impact time τ_s becomes larger and larger, and eventually exceeds τ , the mean time between collisions, so that the collisions overlap. Thus from the very weak collisions there is an essentially continuous perturbation of the atom, and here we expect statistical theory to begin to be valid. Indeed, we saw earlier that these weak collisions produce the line shift, just as would be given by the application of a steady perturbation. Even though Lindholm theory (an impact theory) treats the weak collisions, the calculation is not strictly logically consistent.

(3) Impact theory fails for sufficiently large frequency displacements $\Delta\omega$ from line center, and statistical theory becomes valid. In impact theory, it follows from the general properties of Fourier transforms that the characteristic interruption time τ corresponding to a frequency displacement $\Delta\omega$ is $\tau \sim 1/\Delta\omega$. For sufficiently large $\Delta\omega$'s we will eventually have $\tau \ll \tau_s$, and impact theory breaks down. These values of $\Delta\omega$ correspond to large phase

shifts (i.e., $\Delta\omega\tau_s \gg 1$) and hence to the strong collisions that occur *inside* the Weisskopf radius. It is difficult to construct a theory that makes the transition from impact to statistical theory. A useful conceptualization is to suppose that there is a "boundary" frequency $\Delta\omega_g$ inside of which impact theory holds and beyond which statistical theory is valid. To a fair approximation [see (637; 306)] $\Delta\omega_g \approx \Delta\omega_W$ where $\Delta\omega_W$ denotes the frequency shift produced by a perturber at the Weisskopf radius ρ_W ; i.e.,

$$\Delta\omega_W = (v^p/C_p \psi_p)^{1/(p-1)} \quad (9-78)$$

Note that $\Delta\omega_W$ corresponds to a phase shift of unity. We shall see in §9-4 that equation (9-78) implies that broadening of the hydrogen lines by *ions* follows the *statistical theory*, while *electron* broadening is given by *impact theory*.

(4) Classical impact theory assumes that the collisions are *adiabatic*—i.e., transitions are not induced in the atom. A collision occurring in an impact time τ_s will have Fourier components of frequencies up to $\omega_s \sim 1/\tau_s$. To guarantee that the collision is adiabatic, ω_s should be much smaller than any characteristic transition frequency ω_{ij} ; i.e.,

$$\omega_s = 1/\tau_s \ll \omega_{ij} = |E_i - E_j|/\hbar \quad (9-79)$$

For *nondegenerate levels*, the energy separation is often large enough that the condition stated above will be met. But for *degenerate levels* (e.g., for hydrogen), the energy separation between levels will be proportional to the perturbing field itself; i.e., $|E_i - E_j| \approx \hbar C_p / \rho^p = \hbar \Delta\omega(\rho)$. Then the condition for adiabaticity implies that $\Delta\omega(\rho)\tau_s = \eta(\rho) \gg 1$; that is, *only collisions inside the Weisskopf radius will be adiabatic*. In the case of hydrogen, $\Delta\omega_W$ for ions is very small, and for virtually the entire profile the statistical theory is valid, and the collisions causing the broadening occur inside the Weisskopf radius. Thus the ion broadening will be adiabatic. Precisely the opposite is true for electrons. Here $\Delta\omega_W$ will be large, and almost the whole profile is described by impact theory with the relevant collisions lying outside the Weisskopf radius. The electron broadening is strongly nonadiabatic (and hence must be described by quantum theory). When the adiabatic assumption breaks down, much larger damping parameters than those predicted by classical theory are found; for this reason the modern quantum mechanical results are often drastically different from earlier classical work.

9-4 Collision Broadening: Statistical Theory

The basic picture in this theory is that the atom finds itself radiating in a statistically fluctuating field produced by randomly distributed perturbers. The motion of the perturbers is ignored; this is known as the *quasi-static* approximation. (As we shall see later, this approximation is good for the

slow-moving ions—e.g., protons—in the plasma.) A specific distribution of perturbers produces a definite field; the relative probability of fields of different strengths is thus determined by the statistical frequency with which particle distributions producing the appropriate strengths are realized. For a given value of the field, the oscillation frequency of the radiator is shifted by a definite $\Delta\omega$. The intensity of the radiation at this $\Delta\omega$ is assumed to be proportional to the statistical frequency of occurrence of the appropriate field. Thus the central problem is to determine the probability distribution of the perturbing fields. Once this is known, line profiles can be computed. The applications in this section will be restricted to quasi-static broadening of hydrogen lines by linear Stark-effect interactions with protons (though the theory is relevant in other contexts as well).

THE NEAREST-NEIGHBOR APPROXIMATION

As a first approximation, we assume that the main effect on the radiator results from the *strongest* perturbation acting at any given instant—namely, that from the *nearest neighbor*—and that the effects of all other particles are neglected. Then if $W(r) dr$ is the probability that the nearest neighbor is located on the range $(r, r + dr)$ from the radiator, the frequency spectrum is

$$I(\Delta\omega) d(\Delta\omega) \propto W(r) [dr/d(\Delta\omega)] d(\Delta\omega) \quad (9-80)$$

where it is assumed that $\Delta\omega$ is given by equation (9-53); i.e., $\Delta\omega = C_p/r^p$.

To find $W(r)$ we calculate the probability that a particle is located on the range $(r, r + dr)$, and that none is at a distance less than r . Then assuming a uniform particle density N , $W(r)$ is given by

$$W(r) = \left[1 - \int_0^r W(x) dx \right] (4\pi r^2 N) dr \quad (9-81)$$

where the factor $(4\pi r^2 N) dr$ is the relative probability of a particle lying in the shell $(r, r + dr)$ while the term in square brackets is the probability that no particle lies inside this shell. By differentiation we find

$$\frac{d}{dr} \left[\frac{W(r)}{4\pi r^2 N} \right] = -(4\pi r^2 N) \left[\frac{W(r)}{4\pi r^2 N} \right] \quad (9-82)$$

and thus by integration and normalization,

$$W(r) = 4\pi r^2 N \exp \left(-\frac{4}{3} \pi r^3 N \right) \quad (9-83)$$

It is customary to adopt the *mean interparticle distance* $r_0 = (3/4\pi N)^{1/3}$ as the *reference distance* at which a perturber produces the *normal frequency shift*

$\Delta\omega_0 = C_p/r_0^p$. Then

$$(\Delta\omega/\Delta\omega_0) = (r_0/r)^p \quad (9-84)$$

and equation (9-83) can be rewritten as

$$W(r) dr = \exp[-(\Delta\omega_0/\Delta\omega)^{3/p}] d(\Delta\omega_0/\Delta\omega)^{3/p} \quad (9-85)$$

For the case of *linear Stark effect*, the perturbing field is $F = (e/r^2)$. If we define the *normal field strength* to be

$$F_0 = (e/r_0^2) = e \left(\frac{4}{3} \pi N \right)^{\frac{1}{3}} = 2.5985 e N^{\frac{1}{3}} \quad (9-86)$$

and measure F in units of F_0 [i.e., $\beta \equiv (F/F_0)$], then nearest-neighbor theory yields

$$W(\beta) d\beta = \frac{3}{2} \beta^{-\frac{1}{4}} \exp(-\beta^{-\frac{1}{3}}) d\beta \quad (9-87)$$

Clearly, as $\beta \rightarrow \infty$, $W(\beta) \rightarrow \frac{3}{2}\beta^{-\frac{1}{4}}$; hence statistical theory predicts that, in the wings of a line broadened by linear Stark effect, the profile falls off as $\Delta\omega^{-\frac{1}{4}}$, in contrast with the prediction $\Delta\omega^{-2}$ given by impact theory.

The basic failing of this theory is that the profile is, of course, the result of perturbations by *all* particles, not just the nearest neighbor. To obtain accurate results a more elaborate theory must be constructed.

HOLTSMARK THEORY

The effect of an *ensemble* of particles upon a radiator was determined by Holtsmark (305), who calculated the net vector field, at the position of the radiating atom, from the superposition of the field vectors of all perturbers. An elegant treatment of the problem was given by Chandrasekhar (151); this paper should be consulted for the derivation of the results quoted here.

For an interaction of the form $F = C_p/r^p$, the analysis yields

$$W(\beta) = (2\beta/\pi) \int_0^\infty \exp(-y^{3/p}) y \sin \beta y dy \quad (9-88)$$

Here $\beta = F/F_0$, the normal field strength F_0 now being defined as

$$F_0 = \gamma C_p N^{p/3} \quad (9-89)$$

where $\gamma = \{(2\pi^2 p)/[3(p+3)\Gamma(3/p)\sin(3\pi/2p)]\}^{p/3}$ (9-90)

In particular, for linear Stark effect, $p = 2$, $C_p = e$, and $\gamma = 2.6031$ so that $F_0 = 2.6031 e N^{\frac{1}{3}}$, which differs only inconsequentially from the normal field strength given by nearest-neighbor theory.

The integral in equation (9-88) cannot be evaluated exactly for $p = 2$, and $W(\beta)$ must be expressed in a series expansion. For small β

$$W(\beta) = \left(\frac{4}{3\pi}\right) \sum_{l=0}^{\infty} (-1)^l \Gamma\left(\frac{4l+6}{3}\right) \frac{\beta^{(2l+2)}}{(2l+1)!} \quad (9-91)$$

while for $\beta \gg 1$ one finds an asymptotic expansion

$$W(\beta) = 1.496\beta^{-\frac{1}{2}}(1 + 5.107\beta^{-\frac{1}{2}} + 14.43\beta^{-3} + \dots) \quad (9-92)$$

the leading term of which is essentially the same as given by nearest-neighbor theory. Tabulations of $W(\beta)$ are given in (151) and (629, 28).

DEBYE SHIELDING AND LOWERING OF THE IONIZATION POTENTIAL

In deriving the probability distributions described above, *interactions among the perturbers* were ignored. In reality, the probability that a particle is found in a volume dV is not just $N dV$, but depends also upon the *electrostatic potential* ϕ in dV . For example, if at some point $\phi > 0$, electrons will tend to migrate toward it while ions will tend to migrate away, and vice versa if $\phi < 0$. Following Ecker (203; 204; 205; 206), we may account for these effects schematically by introducing a Boltzmann factor depending on $\psi \equiv (e\phi/kT)$. Thus for electrons and ions, respectively, we write

$$n_e W_e dV = n_e \exp(\psi) dV \approx n_e (1 + \psi) dV \quad (9-93)$$

$$\text{and } n_i W_i dV = n_i \exp(-Z_i \psi) dV \approx n_i (1 - Z_i \psi) dV \quad (9-94)$$

where n_e and n_i are the electron and ion densities, Z_i is the ionic charge, and it is assumed that $\psi \ll 1$. As the plasma is *electrically neutral* over sufficiently large volumes,

$$n_e = \sum_i Z_i n_i \quad (9-95)$$

We now calculate the potential around a particular ion under the simplifying assumption that all particles can be smeared out into an equivalent charge density. We then may use *Poisson's equation*

$$\nabla^2 \phi = -4\pi e \rho \quad (9-96)$$

to determine ϕ , where ρ is given by

$$e\rho = -en_e W_e + e \sum_i Z_i n_i W_i \quad (9-97)$$

In view of equations (9-93) through (9-95), this expression reduces to

$$e\rho = -e\psi \left(n_e + \sum_i Z_i^2 n_i \right) \quad (9-98)$$

Substituting equation (9-98) into (9-96), we may rewrite Poisson's equation as $\nabla^2 \phi = (\phi/D^2)$ where

$$D \equiv (kT/4\pi e^2)^{\frac{1}{2}} \left[n_e + \sum_i Z_i^2 n_i \right]^{-\frac{1}{2}} \quad (9-99)$$

is the *Debye length*. Solving for ϕ we find $\phi = r^{-1}[Ae^{-r/D} + Be^{r/D}]$. Demanding that $\phi \rightarrow 0$ as $r \rightarrow \infty$, we set $B \equiv 0$. Further, to recover the potential of the ion itself as $r \rightarrow 0$, we set $A = Z_i e$. Thus the potential produced by an ion imbedded in a plasma is

$$\phi(r) = [Z_i e \exp(-r/D)]/r \quad (9-100)$$

As is clear from equation (9-100), beyond the Debye length the field of an ion is strongly shielded and rapidly vanishes. Physically this occurs because a charged particle tends to polarize the plasma in its vicinity, and the oppositely charged particles that cluster around it shield the field of the original particle at large distances. Thus the Debye length sets an upper limit on (a) the distance over which two charged particles can effectively interact, (b) the size of a region in which appreciable departures from charge neutrality can occur, and (c) the wavelength of electromagnetic radiation that can propagate through the plasma without dissipating.

In most astrophysical applications we can assume a practically pure hydrogen plasma; then $Z_i = 1$, and $n_i = n_e$, and inserting numerical constants into equation (9-99) we find

$$D = 4.8(T/n_e)^{\frac{1}{2}} \text{ cm} \quad (9-101)$$

Exercise 9-5: Compare Debye lengths in (a) a stellar photosphere with $T = 10^4$ °K, $n_e = 10^{14}$; (b) the solar corona $T = 10^6$ °K, $n_e = 10^8$; (c) an H II region, $T = 10^4$ °K, $n_e = 10^2$.

To calculate the effect of shielding on $W(\beta)$ one can make the very simple assumptions (205; 206) that the field of the perturber is unchanged for $r \leq D$, but is identically zero for $r > D$. One then finds

$$W(\beta, \delta) = (2\beta \delta^{\frac{1}{2}}/\pi) \int_0^\infty e^{-\delta g(y)} y \sin(\delta^{\frac{1}{2}} \beta y) dy \quad (9-102)$$

$$\text{where } g(y) \equiv \frac{3}{2} y^{\frac{1}{2}} \int_y^\infty (1 - z^{-1} \sin z) z^{-\frac{1}{2}} dz \quad (9-103)$$

294 The Line Absorption Profile

and $\delta \equiv \frac{4}{3}\pi D^3 N$, the number of particles contained in the Debye sphere. As $\delta \rightarrow \infty$, one expects to recover the Holtsmark distribution; i.e., $W(\beta, \infty) = W_H(\beta)$. From equation (9-102) one can show that for $\delta \rightarrow \infty$, $W(\beta, \delta) = W_H(\beta) + \delta^{-\frac{1}{2}} F(\beta)$ where $F(\beta)$ is a bounded definite integral. Recovery of the Holtsmark distribution for large δ can also be seen from the asymptotic expansion

$$W(\beta, \delta) = 1.496\beta^{-\frac{1}{2}}(1 + 5.107\beta^{-\frac{1}{2}} - 6.12\delta^{-\frac{1}{2}}\beta^{-2} + \dots) \quad (9-104)$$

which may be compared to equation (9-92). In principle, for small δ the theory should merge continuously into the nearest-neighbor approximation, but in practice, when $\delta \lesssim 5$, the assumptions employed (particularly that a smeared-out charge distribution may be used) break down. A plot of $W(\beta, \delta)$ for several values of δ is shown in Figure 9-1 along with the nearest-neighbor and Holtsmark distributions.

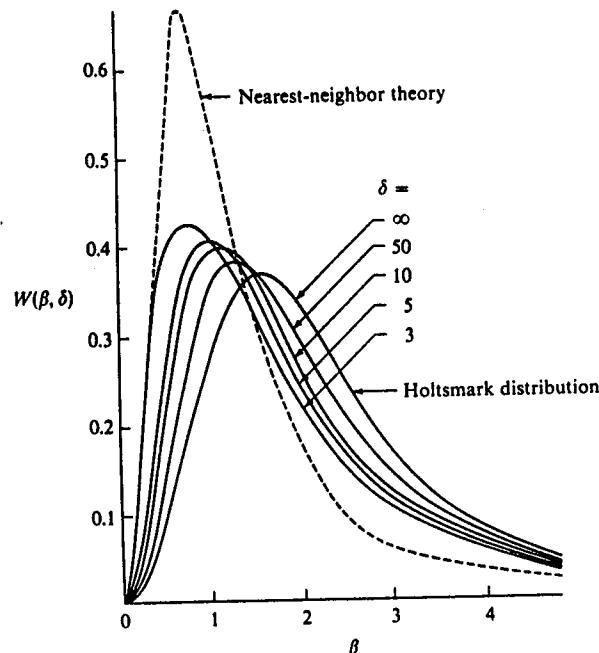


FIGURE 9-1
Probability distribution of field strength at a test point, including shielding effects; δ is the number of charged particles within the Debye sphere. From (205), by permission.

The treatment of perturber interactions given above is somewhat oversimplified. Very precise calculations of the perturber field-strength distributions have been made using cluster-expansion methods (65; 472; 515), and numerical integrations using Monte Carlo techniques (308; 309; 310; 482). Most modern treatments of hydrogen-line broadening use these refined distribution functions. In practice, the effects of shielding are often quite important in laboratory plasmas, while in stellar atmospheres the densities are so low that the number of particles in a Debye sphere is large ($\delta \gtrsim 100$), and the departures from the Holtsmark distribution are not large.

The presence of nearby charges partly neutralizes the effects of the nuclear charge upon an orbital electron and thereby weakens the potential well in which the electron is bound. The reduction in binding energy can be calculated, using the Debye potential in equation (9-100), as

$$\Delta E = (Ze^2/r)[\exp(-r/D) - 1] \approx -Ze^2/D \quad (9-105)$$

for $r \ll D$. An electron in a state that, in the unperturbed atom, lies at an energy $\Delta\chi$ below the ionization limit, can be considered unbound if $\Delta\chi \lesssim \Delta E$. That is, the ionization potential of the atom is decreased by an amount

$$\Delta\chi = Ze^2/D = (27.2Za_0/D) \text{ eV} = 3 \times 10^{-8} Z n_e^{-\frac{1}{2}} T^{-\frac{1}{2}} \text{ eV} \quad (9-106)$$

where use has been made of equation (9-101). This calculation of $\Delta\chi$ is only schematic. A discussion from several points of view of the lowering of ionization potentials in a plasma is given in (178).

THE QUASI-STATIC ION BROADENING OF HYDROGEN LINES

In the absence of a perturbing field, each level of hydrogen is degenerate with $2n^2$ sublevels. Analyses by K. Schwarzschild (563) and Epstein (208) showed that, when an electric field is applied, these sublevels separate and, because hydrogen has a permanent dipole moment, the energy shift is directly proportional to the applied field strength F (linear Stark effect). If no other broadening mechanisms are operative, the line profile will consist of a number of Stark components, arising from transitions between the sublevels of the lower and upper states. Each Stark component has a characteristic relative intensity I_k (561) and will be displaced from line center by a characteristic shift

$$\Delta\lambda_k = (3h\lambda^2 n_k / 8\pi^2 c meZ)F = C_k F \quad (9-107)$$

where Z is the charge on the atom ($= 1$ for hydrogen) and n_k is an integer. The line that we observe will be a superposition of these components, weighted by their relative intensities and the probability of being shifted to the appropriate wave-length position.

The Stark pattern of a hydrogen line is symmetric about line center with identical components at $\pm k$ with $I_{-k} = I_k$, $C_{-k} = C_k$ [see, e.g., (638, 320; 629, 73)]. Assuming that the intensities are normalized such that $\sum_k I_k = 1$ (where the sum extends over all components), then the line profile will be

$$\begin{aligned} I(\Delta\lambda) d(\Delta\lambda) &= \sum_k I_k W(F/F_0)(dF/F_0) \\ &= \sum_k I_k W(\Delta\lambda/C_k F_0) d(\Delta\lambda)/C_k F_0 \end{aligned} \quad (9-108)$$

It is customary to define the parameter α as

$$\alpha \equiv \Delta\lambda/F_0 \quad (9-109)$$

where F_0 is the normal field strength $F_0 = 2.60eN^3$. Then the line profile is given by a function $S(\alpha)$,

$$S(\alpha) d\alpha = \sum_k I_k W(\alpha/C_k)(d\alpha/C_k) \quad (9-110)$$

which is normalized on the range $(-\infty, \infty)$ for α . The absorption cross-section per atom can be written as

$$\alpha_v(\Delta\lambda) = (\pi e^2/mc) f S(\Delta\lambda/F_0)(\lambda^2/cF_0) \quad (9-111)$$

Extensive tables of C_k , I_k , and $S(\alpha)$ for numerous hydrogen lines are given in (634). The largest C_k for a line of upper quantum number n increases as n^2 , and recalling that for $\beta \gg 1$, $W(\beta) \propto \beta^{-1}$, we see from equation (9-108) or (9-110) that the Stark widths of lines rise rapidly up a series, roughly as n^3 .

Let us now consider when the quasi-static profiles are applicable. Let $\bar{n}_k = \sum I_k n_k / \sum I_k$, the sum being taken over positive values only. Then, writing $\Delta\omega = C_2/r^2$ and $F = e/r^2$, it follows from equation (9-107) that

$$C_2 = (3h\bar{n}_k/4\pi m) = 1.738\bar{n}_k \quad (9-112)$$

with $\bar{n}_k \sim \frac{1}{2}n(n-1)$ for $n \gg 1$. From equation (9-78) the wavelength shift delimiting the transition between the impact and statistical theories is (for $p = 2$)

$$\Delta\lambda_w = (v^2\lambda^2/2\pi^3 c C_2) \quad (9-113)$$

Notice that $\Delta\lambda_w \propto v^2$, and thus $\Delta\lambda_w(\text{electron}) \sim 10^3 \Delta\lambda_w(\text{proton})$. Using equations (9-112) and (9-50), we obtain the results listed in Table 9-2. It is obvious that the ion broadening is very well described by the quasi-static theory (especially when we note that Doppler motions will dominate in the core). The electrons, however, are in the impact-broadening regime except

TABLE 9-2
Transition Wavelength $\Delta\lambda_w(\text{\AA})$ between
Statistical and Impact Broadening for
Hydrogen Lines

Line	Perturber	T	
		$2.5 \times 10^4 \text{ }^\circ\text{K}$	$10^4 \text{ }^\circ\text{K}$
$H\alpha$	Electrons	580.0	230.0
	Protons	0.63	0.25
$H\beta$	Electrons	120.0	48.0
	Protons	0.13	0.05
$H\gamma$	Electrons	48.0	19.0
	Protons	0.05	0.02
$H\delta$	Electrons	32.0	13.0
	Protons	0.03	0.01

at very large displacements from line center and, as mentioned in §9-3, are nonadiabatic. The complete profile consists of the effects of both ions and electrons and, as we shall see in §9-5, the latter increase the line-broadening markedly. The functions $S(\alpha)$ for ions alone seriously underestimate the hydrogen-line widths, and a satisfactory theoretical description of stellar hydrogen-line profiles became possible only after the development of the quantum mechanical line-broadening theory.

9-5 Quantum Theory of Line Broadening

Quantum mechanical calculations yield precise profiles for pressure-broadened lines. The development of this theory brought about a major improvement in one of the most important (and difficult) applications of atomic physics in the analysis of stellar spectra. Excellent discussions of the general theory can be found in (62; 63; 64; 73, Chap. 13; 179; 264, Chap. 4; 268; 582); only a brief outline will be presented here. We shall focus attention on the case where the atom suffers impact broadening from electrons and quasi-static broadening by the ions.

THE LINE PROFILE

As was shown in Chapter 4, the power radiated by an *isolated atom* in a transition from an upper state j to a lower state i is [cf. equation (4-62)]

$$P = (4\omega^4/3c^3)|\langle i|\mathbf{d}|j\rangle|^2 \quad (9-114)$$

where \mathbf{d} is the atomic dipole moment. The total emission, summed over all possible substates contributing to a line is

$$P(\omega) = (4\omega^4/3c^3) \sum_{i,j} \rho_j \delta(\omega - \omega_{ij}) |\langle i|\mathbf{d}|j\rangle|^2 \quad (9-115)$$

Here ρ_j is the probability that an atom is in the upper state j ; in thermodynamic equilibrium

$$\rho_j = \langle j|\rho|j\rangle = \exp(-E_j/kT)/U(T) \quad (9-116)$$

where $U(T)$ is the partition function.

To calculate the broadening of lines emitted by an *atom in a plasma*, we consider the radiating system to consist of atom *plus* perturbers, and generalize the meaning of states $|i\rangle$ and $|j\rangle$ to include perturbers also. The *profile* of the line is then written as

$$I(\omega) = \sum_{i,j} \rho_j \delta(\omega - \omega_{ij}) |\langle i|\mathbf{d}|j\rangle|^2 \quad (9-117)$$

Now ρ_j refers to the probability of a particular state of atom *and* perturber. If the plasma is in thermal equilibrium, ρ_j is proportional to

$$e^{-H/kT} = e^{-(H_A + H_P + V)/kT} \quad (9-118)$$

where H is the total Hamiltonian of the system, H_A and H_P are the Hamiltonians of the atom and perturber alone, respectively, and V is the interaction Hamiltonian.

It is easiest to account for collisions using the Fourier transform

$$\phi(t) = \int_{-\infty}^{\infty} I(\omega) e^{i\omega t} d\omega = \sum_{i,j} \rho_j e^{i\omega_{ij} t} |\langle i|\mathbf{d}|j\rangle|^2 \quad (9-119)$$

which is analogous to the classical *autocorrelation function* [see (72, 498)]. The effects of collisions at a specific impact parameter upon the autocorrelation function (and statistical averages over all possible perturber paths) can be calculated directly. The intensity profile then follows from the inverse Fourier transform [cf. equation (9-7)].

The problem at hand is to obtain an expression for $\phi(t)$. To do this we must find the change with time of the eigenstates $|i\rangle$ and $|j\rangle$ under the effects of perturbing collisions in terms of a *time-development operator* $T(t, 0)$. This operator is defined such that a state of the system at time t is related to the state at time $t = 0$ by

$$|\alpha, t\rangle = T(t, 0)|\alpha, 0\rangle \quad (9-120)$$

Now $|\alpha, t\rangle$ satisfies the Schrödinger equation

$$H|\alpha, t\rangle = i\hbar(d|\alpha, t\rangle/dt) \quad (9-121)$$

Substituting equation (9-120) into (9-121) we can derive a Schrödinger equation for $T(t, 0)$. Noting that $|\alpha, 0\rangle$ is *fixed* in time, we find

$$HT(t, 0) = i\hbar[dT(t, 0)/dt] \quad (9-122)$$

which has the solution

$$T(t, 0) = e^{-iHt/\hbar} \quad (9-123)$$

where the exponential is to be understood as an *operator*. We may now rewrite equation (9-119) in terms of time-development operators, *including perturbations*; we have

$$\begin{aligned} \phi(t) &= \sum_{i,j} \rho_j e^{i(E_j - E_i)t/\hbar} |\langle i|\mathbf{d}|j\rangle|^2 \\ &= \sum_{i,j} \rho_j \langle j|\mathbf{d}|i\rangle e^{-iE_it/\hbar} \langle i|\mathbf{d}|j\rangle e^{iE_jt/\hbar} \\ &= \sum_{i,j} \rho_j \langle j|dT|i\rangle \langle i|\mathbf{d}T^\dagger|j\rangle \end{aligned} \quad (9-124)$$

But the *expansion rule*, relative to a *complete set of states* $|\gamma\rangle$, is

$$\langle \alpha|\beta\rangle = \sum_{\gamma} \langle \alpha|\gamma\rangle \langle \gamma|\beta\rangle \quad (9-125)$$

so we see that equation (9-124) can be rewritten as

$$\phi(t) = \sum_j \rho_j \langle j|dT dT^\dagger|j\rangle = \text{tr}(\rho dT dT^\dagger) \quad (9-126)$$

The trace includes both atomic and perturber states. This expression is quite general; the detailed form of ϕ depends upon the form of T .

THE CLASSICAL PATH APPROXIMATION

Let us now consider the calculation of T in more detail. Let the wave function ψ describing the solution of the complete system of atom plus perturber be the solution of the equation

$$i\hbar(d\psi/dt) = (H_A + H_P + V)\psi \quad (9-127)$$

Note that H_A is independent of perturber coordinates, H_P is independent of atomic coordinates, while V depends on both. Further, because the atom

is in a static ionic field of strength F , $H_A = H_A(F)$. For brevity, F will be suppressed here but included explicitly in the final result.

To make progress, we now assume that the perturber follows its *classical path*—i.e., a *straight line* past a neutral atom and a *hyperbola* around an ion. This assumption will be valid when the *deBroglie wavelength* is small compared to the impact parameter for those collisions that dominate the broadening. That is, we must have $\rho \gg \lambda = (\hbar/mv)$ or $mvp \gg \hbar$. But mvp is just the angular momentum of the perturber ($=\hbar l$); the criterion just stated is equivalent to the requirement that the quantum number l be much larger than unity. Under such circumstances the classical-particle picture would be expected to be valid on the basis of the *correspondence principle*. The validity of the approximation must always be checked in line-broadening calculations; in most cases of astrophysical interest it is found to hold. Some perturbers that violate the condition can always be expected, but the approximation remains useful if they do not dominate the broadening [see also (73, 498 ff.)]. A theory that does not use the classical path approximation is given in (64).

The wave function for the system of atom plus perturber is assumed to be *separable*; i.e., $\psi(t) = \alpha(t)\pi(t)$, where $\alpha(t)$ and $\pi(t)$ are the atomic and perturber wave functions. We further suppose the perturber path to be fixed and to be independent of the state of the atom with which the interaction takes place. In this way the effect of the perturber on the atom is taken into account, but the back-reaction of the atom on the perturber is ignored. This will be valid if the energy gained or lost by the perturber (of order $\hbar\Gamma$, where Γ is the linewidth) is much less than its kinetic energy (kT), a condition that is almost always satisfied. A few collisions will always occur in which large energy exchanges take place, but again these will not invalidate the assumption if they do not dominate the broadening.

Under the assumptions made above, $\pi(t)$ is the solution of

$$i\hbar[d\pi(t)/dt] = H_p\pi(t) \quad (9-128)$$

and the time-development operator for the perturber is

$$T_p(t, 0) = e^{-iH_p t/\hbar} \quad (9-129)$$

Now consider the Schrödinger equation for the atom alone. To obtain it, we multiply equations (9-127) and (9-128) by π^* , integrate over perturber coordinates, and subtract to find

$$i\hbar[d\alpha(t)/dt] = \left(H_A + \int \pi^* V \pi d\tau_p \right) \alpha(t) \quad (9-130)$$

If the perturber wave packets are indeed narrow enough that the perturbers can be considered to be classical particles on classical paths, then we can

make the identification

$$\int (\pi^* V \pi) d\tau_p \rightarrow V_{cl}(t) \quad (9-131)$$

where $V_{cl}(t)$ is the *classical interaction potential*; this is the essence of the *classical path approximation*. The Schrödinger equation for the atom's time-development operator then becomes

$$i\hbar[dT_A(t, 0)/dt] = [H_A + V_{cl}(t)]T_A(t, 0) \quad (9-132)$$

and the time-development operator for the complete system is

$$T(t, 0) = T_A(t, 0)T_p(t, 0) = T_A(t, 0)e^{-iH_p t/\hbar} \quad (9-133)$$

Finally, we write the probability-density matrix ρ as $\rho = \rho_A \rho_p$ where ρ_A refers to atomic states only, and ρ_p refers to perturber states only and is diagonal in the perturber coordinates. Again, this may be done if the back-reaction of the atom on the perturber can be neglected.

When these expressions for ρ and T are inserted into equation (9-126) for $\phi(t)$, and the separated form of ψ is recalled, we find

$$\phi(t) = \text{tr}\{\rho_A dT_A dT_A^\dagger\} \quad (9-134)$$

The trace over perturber states has reduced merely to a *thermal average over all perturbers* (denoted by braces) and trace is now carried out over *atomic states only*.

THE IMPACT APPROXIMATION

To calculate $\phi(t)$ as given by equation (9-134) we assume that both the initial state a and the final state b consist of several substates, denoted by α and β , respectively, and that dipole transitions exist only between substates of a and b , but that radiative transitions among the substates of a or b can be ignored. On the other hand, we ignore collision-induced transitions between states a and b , and assume that collisions can result only in transitions among substates of a or b . That is, we have $\langle \alpha | d | \alpha' \rangle = 0$, $\langle \beta | d | \beta' \rangle = 0$, and $\langle \alpha | T_{a,b} | \beta \rangle = 0$. Then writing out the trace in equation (9-134) we find

$$\phi(t) = \rho_a \sum_{\alpha, \alpha', \beta, \beta'} \langle \alpha | d | \beta \rangle \langle \beta' | d | \alpha' \rangle \langle \alpha | \langle \beta | \{ T_b T_a^* \} | \alpha' \rangle | \beta' \rangle \quad (9-135)$$

Here we have neglected the variation of ρ_a among the upper substates and have noted that only the time-development operators can depend upon the statistical averages.

It is convenient to replace the complete time-development operator T by U , the time-development operator in the *interaction representation*, defined as

$$U(t, 0) = e^{iH_{\text{A}}t/\hbar} T(t, 0) \quad (9-136)$$

This definition factors out the time-behavior of an eigenstate and isolates the perturbation effects; for an unperturbed eigenstate $U(t, 0) \equiv 1$. By substitution into equation (9-135) we have

$$\phi(t) = \rho_a \sum_{\alpha, \alpha', \beta, \beta'} \langle \alpha | d | \beta \rangle \langle \beta' | d | \alpha' \rangle \langle \alpha | \langle \beta | e^{-iH_b t/\hbar} \{ U_b U_a^* \} e^{iH_a t/\hbar} | \alpha' \rangle | \beta' \rangle \quad (9-137)$$

Substituting equation (9-136) into (9-132) we obtain a Schrödinger equation for $U(t, 0)$,

$$ih[dU(t, 0)/dt] = e^{iH_{\text{A}}t/\hbar} V_{cl}(t) e^{-iH_{\text{A}}t/\hbar} U(t, 0) \equiv V'_{cl}(t) U(t, 0) \quad (9-138)$$

We solve this equation by *iteration*, starting with $U(t, 0) = 1$ as a first approximation, to obtain a solution of the form

$$U(t, 0) = 1 + (ih)^{-1} \int_0^t V'_{cl}(t_1) dt_1 + (ih)^{-2} \int_0^t dt_2 V'_{cl}(t_2) \int_0^{t_2} dt_1 V'_{cl}(t_1) + \dots \quad (9-139)$$

Now consider the calculation of the statistical average $\{U_b U_a^*\}$. We proceed in a way analogous to that used in Lindholm theory. That is, we write the change in $\{U_b U_a^*\}$, caused by a specific collision (treated as an *impact*), in some time Δt as $\Delta\{U_b U_a^*\} = \{U_b(t + \Delta t, t) U_a^*(t + \Delta t, t) - 1\} \{U_b U_a^*\}$, where again we argue that the changes in $(t, t + \Delta t)$ are statistically independent of the current values, so that we can replace the average of the product by the product of the averages. Then using equation (9-139) on the interval $(t, t + \Delta t)$ one obtains an explicit expression for $\{U_b(t + \Delta t, t) U_a^*(t + \Delta t, t) - 1\}$ [see, e.g., (264, 70; 268, 37)] of the form

$$\{U_b(t + \Delta t) U_a^*(t + \Delta t) - 1\} = e^{i(H_b - H_a)t/\hbar} (\Phi_{ab} \Delta t) e^{-i(H_b - H_a)t/\hbar} \quad (9-140)$$

We thus obtain a differential equation for $\{U_b U_a^*\}$, namely

$$\Delta\{U_b U_a^*\} = e^{i(H_b - H_a)t/\hbar} \Phi_{ab} e^{-i(H_b - H_a)t/\hbar} \{U_b U_a^*\} \Delta t \quad (9-141)$$

whose solution is

$$\{U_b(t, 0) U_a^*(t, 0)\} = e^{i(H_b - H_a)t/\hbar} \exp[-i(H_b - H_a)t/\hbar + \Phi_{ab}t] \quad (9-142)$$

Now, substituting into equation (9-137), we obtain

$$\phi(t) = \rho_a \sum_{\alpha, \alpha', \beta, \beta'} \langle \alpha | d | \beta \rangle \langle \beta' | d | \alpha' \rangle \langle \alpha | \langle \beta | e^{i(H_b - H_a)t/\hbar + \Phi_{ab}t} | \alpha' \rangle | \beta' \rangle \quad (9-143)$$

Then performing the inverse Fourier transformation, and reintroducing the quasi-static ion field F , we obtain the intensity profile

$$I(\omega, F) = \frac{\rho_a}{\pi} \Re e \left(\sum_{\alpha, \alpha', \beta, \beta'} \langle \alpha | d | \beta \rangle \langle \beta' | d | \alpha' \rangle \times \langle \alpha | \langle \beta | \left[i\omega - \Phi_{ab}(F) - \frac{i}{\hbar} [H_a(F) - H_b(F)] \right]^{-1} | \alpha' \rangle | \beta' \rangle \right) \quad (9-144)$$

where use has been made of the fact that $\Phi_{ab}(F)$ is found to have a negative real part. If $W(F)$ is the probability of an ion field of strength F , the final profile, averaged over all ion fields is

$$I(\omega) = \frac{\rho_a}{\pi} \int_0^\infty W(F) \Re e \left(\sum_{\alpha, \alpha', \beta, \beta'} \langle \alpha | d | \beta \rangle \langle \beta' | d | \alpha' \rangle \times \langle \alpha | \langle \beta | \left[i\omega - \Phi_{ab}(F) - \frac{i}{\hbar} [H_a(F) - H_b(F)] \right]^{-1} | \alpha' \rangle | \beta' \rangle \right) dF \quad (9-145)$$

Equation (9-145) is quite general, and has been used in most quantum mechanical calculations of Stark-broadened line profiles. The result is valid if (1) the interval Δt in equation (9-140) can be chosen to include a complete collision; (2) when the collisions overlap, they are weak enough that their contributions to the iterative solution for U are simply additive; and (3) the perturbers can be treated as classical particles. These validity criteria must be checked in each case.

APPLICATION TO HYDROGEN

One of the most important applications of the theory outlined above has been to the calculation of the effects of electron impacts upon the broadening of hydrogen lines. The theory has reached a very refined state, and the theoretically predicted profiles are in excellent agreement with laboratory measurements (671) and provide satisfying fits to observed stellar profiles (see Figure 10-4).

The impact theory outlined above has been intensively applied by Griem and his collaborators (263; 265; 270; 271). Although the evaluation of Φ_{ab} is straightforward in principle, it is complicated in practice, for it entails cutoffs both at small and large impact parameters, the former arising from strong collisions inside the Weisskopf radius (which are no longer correctly described by the iterative series development for U , the time-development operator) and the latter to account for Debye shielding effects. Further cutoff procedures are required to allow for the transition of the electrons from the impact- to statistical-broadening regimes. This work culminated in the publication of extensive tables (353; 356) for the first four members of the Lyman

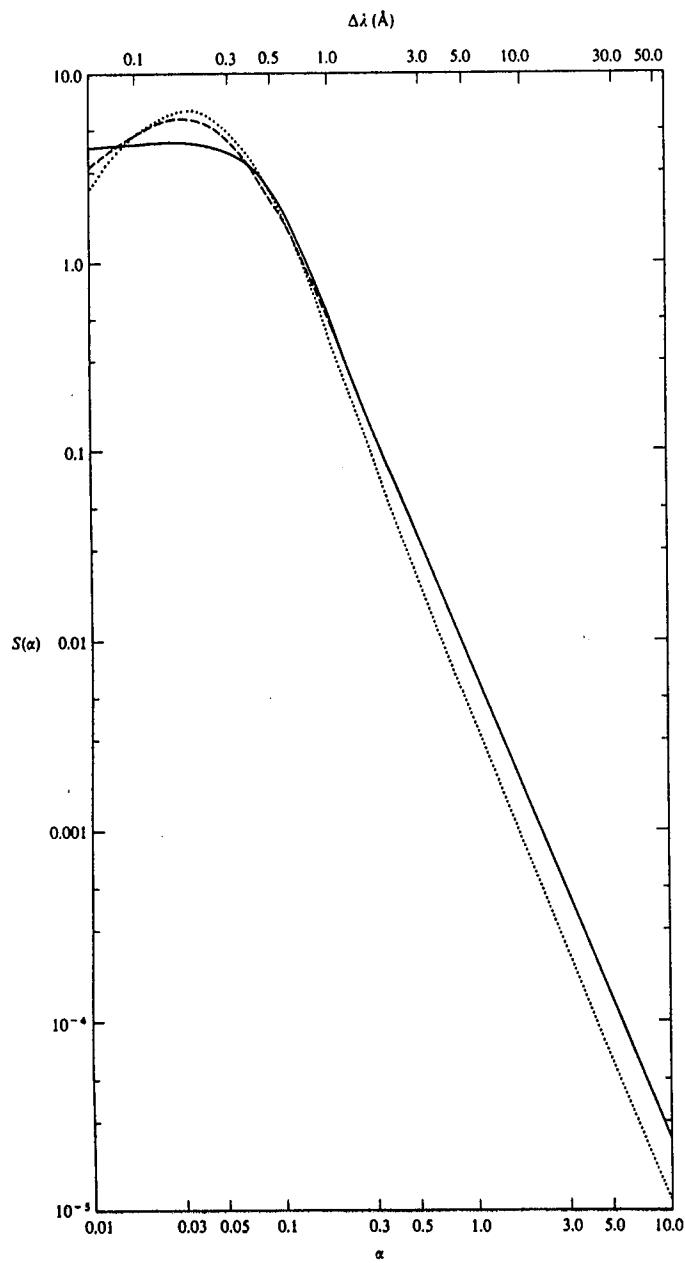


FIGURE 9-2
Stark profiles for $H\alpha$ at $n_e = 3.16 \times 10^{14} \text{ cm}^{-3}$, $T = 10^4 \text{ }^\circ\text{K}$. Dotted curve: Holtsmark profile for ion broadening only (634). Dashed curve: Profile for quasi-static ion broadening and impact broadening by electrons (650). Solid curve: Electron and ion-broadened profile

and Balmer series lines. These tables give $S(\alpha)$, analogous to the profile described by equations (9-109) through (9-111), but *including the effects of electron impacts*. An alternative approach developed by Cooper, Smith, and Vidal (581; 647; 648; 649) uses a *unified theory* that accounts automatically for the transition of electrons from the impact to quasi-static regimes. Extensive tables of results based on this theory have been published (650) for the first four members of the Lyman and Balmer series in temperature-density ranges appropriate to stellar atmospheres. These tables include the convolution of the Stark profile with a thermal velocity distribution of the hydrogen atoms. Figure 9-2 compares the profile of $H\delta$ from the unified theory (including both electron and ion broadening) with the quasi-static theory for ions only. For higher series members one may use an approximate theory (262), after corrections to some of the matrix elements are made (265) [see also (45)].

In astrophysical applications, the hydrogen lines are significantly affected by broadening mechanisms other than the Stark effect. The core of the line is dominated by Doppler broadening. The effects of radiation and resonance damping may be important in the wings at low electron densities. Assuming these mechanisms are all uncorrelated, we may account for their combined effects by a convolution procedure. Folding the Doppler profile with the Lorentz profiles from radiation and resonance damping gives a Voigt profile, $H(a, v)$, where $a = (\Gamma_{\text{rad}} + \Gamma_{\text{res}})/4\pi \Delta v_D$, and $v = (\Delta v/\Delta v_D)$. This Voigt profile is then convolved (numerically) with the Stark profile $S(\alpha)$, yielding the cross-section

$$\alpha_v(\Delta v) = (\pi^4 e^2 / mc) \int_{-\infty}^{\infty} S^*(\Delta v + v \Delta v_D) H(a, v) dv \quad (9-146)$$

where S^* is $S(\alpha)$ converted to frequency units.

HYDROGENIC IONS

Hydrogenic ions (of charge Z) have Stark patterns identical to those of hydrogen, though the energies are, of course, different. The profile from ion-broadening alone is

$$S(\alpha) = Z^5 S_{\text{QS}}(Z^5 \alpha) \quad (9-147)$$

where S_{QS} is the quasi-static hydrogen profile given in (634). Note that on a wavelength scale the ion lines are narrower by a factor of $1/Z^5$. Indeed, because the Stark widths decrease for ions while the radiative transition probabilities increase, there comes a point where Stark broadening can be neglected compared to Γ_{rad} .

The effects of electron broadening for hydrogenic ions are similar to those for hydrogen, though the expression for Φ_{ab} changes because now the

perturber moves on a *hyperbola* around the positively charged ion instead of a *straight-line* path. Early calculations of the broadening of the lines He II $\lambda\lambda 3203, \lambda 4686$ (prominent in O-star spectra) are given in (272), and much-improved calculations for He II $\lambda\lambda 256, 304, 1085, 1216, 1640, 4686$, and 3203 are given in (354; 355). Unified theory (258; 260) calculations for He II $\lambda 304$ are given in (259). Unfortunately, precise calculations for the astrophysically important Pickering series lines ($4 \rightarrow n$) (e.g., $\lambda\lambda 10124, 5412, 4542, 4200$, etc.) are not yet available, and only an approximate theory exists (262; 265; 45).

NEUTRAL HELIUM LINES

The lines of He I are prominent in the spectra of B-stars. Here the electron impact broadening and quasi-static ion broadening act by quadratic Stark effect, and for isolated lines yield profiles of the form

$$I(\Delta\omega) = \left(\frac{w}{\pi}\right) \int_0^\infty \frac{W(F) dF}{(\Delta\omega - d - C_4 F^2/e^2)^2 + w^2} \quad (9-148)$$

where $W(F)$ is the probability of a field of strength F , w is the electron-impact *width* of the line, and d is the line *shift*. Note that because F enters as F^2 , the ionic fields always skew the line components in one direction and thus lead to an *asymmetric* profile.

Explicit expressions for w and d are given in (269; 264, 81–86; 268, §II.3cα; 180) for *isolated* lines; tables allowing the calculation of the profile in terms of convenient dimensionless units are given in (269). Detailed numerical results for w and d for several lines are given in (269), and improved results are given in (180) and (67). A much more interesting (and difficult) case is presented by the diffuse series lines (2^3P-n^3D), (2^1P-n^1D) for $n \geq 4$. As was first recognized by Struve (614; 615), the (2^3P-4^3D) He I $\lambda 4471$ line shows a “*forbidden*” component (2^3P-4^3F) at $\lambda 4470$. The other diffuse series lines also show these components which arise from the mixing of the (3D , 3F) or (1D , 1F) states in the presence of the electric fields in the plasma. As the diffuse-series lines are among the best observed in stellar spectra, a reliable theory for them is of great interest. The first attempts at constructing such broadening theories (66; 245; 266) were not too successful, for they gave “*forbidden*” components that were too narrow and too intense. Compared to observed stellar profiles the theoretical predictions showed too much contrast between the “*forbidden*” absorption component and the gap between it and the allowed component; the theory also disagreed with experimental measurements (122; 123; 124). A comprehensive new theory was then developed for the He I $\lambda 4471$ (68) and $\lambda 4921$ (2^1P-4^1D) lines (69); predictions based on this theory are in excellent agreement with observed stellar spectra (438; 439).

OTHER LIGHT ELEMENTS

Electron collisions broaden the lines of other elements observed in stellar spectra. Widths and shifts for these lines may be calculated using techniques similar to those employed for He I [see (268, §§II.3c, II.3d)]. Extensive results for neutral atoms are given in (264, 454–527; 268, Appendix IV; 84). For charged ions, the Coulomb interaction between radiator and perturber implies a hyperbolic path for the latter (110; 111), and line-widths calculated allowing for this are substantially larger than those calculated with a straight-line path. Extensive results of detailed calculations for ions are given in (268, Appendix V; 111; 548; 167; 549; 550). Convenient approximate formulae for estimating Stark-widths are given in (267; 183; 549).