where w_0 is an appropriate weight factor, and consider the transfer problem only at the continuum head, $\nu = \nu_0$ [see also (**195**)]. If we define $\varepsilon \equiv C/[4\pi w_0(\alpha_0/h\nu_0)B_0 + C]$ then

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

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

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

Assuming $\tilde{\varepsilon}$ 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

$$J_0 = B_0\{\tilde{\varepsilon}^{\frac{1}{2}} + 1 - \exp[-\sqrt{3\tilde{\varepsilon}}\,\tau_0]\}/(1 + \tilde{\varepsilon}^{\frac{1}{2}}) \quad (7\text{-}134)$$

which shows that (a) the solution *thermalizes* only at depths $\gtrsim 1/\tilde{\varepsilon}^{\frac{1}{2}}$, and (b) the departure coefficient at the surface is $b_1(0) \approx \tilde{\varepsilon}^{-\frac{1}{2}}$. Because $\tilde{\varepsilon} \ll 1$, *the departure from LTE is large*, and *persists to great depths*. The actual value of $\tilde{\varepsilon}$ 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 $T \sim 25{,}000°\text{K}$, $r \sim 10^{-2}$ [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

$$d^2(f_\nu J_\nu)/d\tau_\nu^2 = J_\nu - (\eta_\nu + n_e\sigma_e J_\nu)/\chi_\nu = J_\nu - S_\nu \quad (7\text{-}135)$$

where (omitting bound–bound transitions)

$$\chi_\nu = \sum_i (n_i - n_i^* e^{-h\nu/kT})\alpha_{i\kappa}(\nu) + \sum_\kappa n_e n_\kappa \alpha_{\kappa\kappa}(\nu, T)(1 - e^{-h\nu/kT}) + n_e\sigma_e \quad (7\text{-}136)$$

and

$$\eta_\nu = (2h\nu^3/c^2)e^{-h\nu/kT}\left[\sum_i n_i^*\alpha_{i\kappa}(\nu) + \sum_\kappa n_e n_\kappa \alpha_{\kappa\kappa}(\nu, T)\right]$$

$$= B_\nu\left[\sum_i n_i^*\alpha_{i\kappa}(\nu) + \sum_\kappa n_e n_\kappa \alpha_{\kappa\kappa}(\nu, T)\right](1 - e^{-h\nu/kT}) = \kappa_\nu^* B_\nu \quad (7\text{-}137)$$

We now divide the spectrum into a series of characteristic frequency ranges ($\nu_{i0} \leqslant \nu \leqslant \nu_{i1}$) upon which we assume the opacity and emissivity of a particular continuum $i \to \kappa$ is much larger than all others. For example we might assume that the Lyman continuum dominates for all frequencies such that $\lambda \leqslant 912$ Å, the Balmer continuum dominates for 912 Å $\leqslant \lambda \leqslant 3650$ Å, the free–free continuum dominates for $\lambda \geqslant \lambda_I$, 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 ($\nu_{i0} \leqslant \nu \leqslant \nu_{i1}$),

$$(\eta_\nu/\chi_\nu) = [n_i^*/(n_i - n_i^* e^{-h\nu/kT})]\xi_{i\nu}B_\nu = \xi_{i\nu}B_\nu/(b_i - e^{-h\nu/kT}) \quad (7\text{-}138)$$

where the three quantities

$$\xi_{i\nu} \equiv [(n_i - n_i^* e^{-h\nu/kT})/n_i^*](\kappa_\nu^*/\chi_\nu) \quad (7\text{-}139)$$

f_ν, and τ_ν are all evaluated using *current values* of the temperature and level populations, whereas the term $(b_i - e^{-h\nu/kT})$ 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)]

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

where $R_{i\kappa}$ and $R_{\kappa i}$ are defined, as usual, by equations (5-66) and (5-67). Solving for $b_i - e^{-h\nu/kT}$ we then have

$$(b_i - e^{h\nu/kT})^{-1} = \left(R_{i\kappa} + \sum_{j\neq i}^{\kappa} C_{ij}\right)\Bigg/\Bigg[R_{\kappa i} - e^{-h\nu/kT}R_{i\kappa} + \sum_{j\neq i}^{I}(b_j - e^{-h\nu/kT})C_{ij} + C_{i\kappa}(1 - e^{-h\nu/kT})\Bigg] \quad (7\text{-}141)$$

so that, defining $\zeta_\nu \equiv n_e\sigma_e/\chi_\nu$ (again using current values), the source function in equation (7-135) on the range $(\nu_{i0} \leqslant \nu \leqslant \nu_{i1})$ can be written as

$$S_\nu = \xi_{i\nu}\left(\gamma_{i\nu}\int_{\nu_{i0}}^{\nu_{i1}} \Phi_{i\nu}J_\nu\, d\nu + \varepsilon_{i\nu}B_\nu\right) + \zeta_\nu J_\nu \tag{7-142}$$

where

$$\Phi_{i\nu} \equiv 4\pi\alpha_{i\kappa}(\nu)/h\nu$$

$$\gamma_{i\nu} \equiv B_\nu/D_{i\nu}$$

$$\varepsilon_{i\nu} \equiv \left(\int_{\nu_{i1}}^{\infty} \Phi_{i\nu}J_\nu\, d\nu + \sum_{j\neq i}^{\kappa} C_{ij}\right)\Big/ D_{i\nu} \tag{7-143}$$

and $D_{i\nu}$ is the term in square brackets in the denominator of equation (7-141). In the evaluation of $\gamma_{i\nu}$ and $\varepsilon_{i\nu}$ we again use current values for the b_j's of *other* levels and for the terms in J_ν and T; it is thus clear that a large amount of information is *lagged* in this technique. Note that $\varepsilon_{i\nu}$ contains an *overlap integral* from those frequency ranges outside the presumed range of dominance (ν_{i0}, ν_{i1}) of the transition $i \to \kappa$; 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

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

on the range $(k_{i0} \leqslant k \leqslant k_{i1})$ delineated for the ith transition.

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

$$\int_0^\infty \chi_\nu(J_\nu - S_\nu)\, d\nu = \sum_{k=1}^{K} w_k\chi_kJ_k - \sum_{i=1}^{I}\sum_{k=k_{i0}}^{k_{i1}} w_k\chi_k \times\left[\xi_{ik}\left(\gamma_{ik}\sum_{l=k_{i0}}^{k_{i1}} w_l\Phi_{il}J_l + \varepsilon_{ik}B_k\right) + \zeta_kJ_k\right] = 0 \tag{7-145}$$

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

$$\Delta T = \sum_{k=1}^{K} w_k\psi_kJ_k - E_1 \tag{7-146}$$

where

$$\psi_k \equiv \left[\chi_k(1-\zeta_k) - \Phi_{ik}\sum_{l=k_{i0}}^{k_{i1}} w_l\chi_l\xi_{il}\gamma_{il}\right]\Big/ E_2, \quad \begin{matrix}(k_{i0} \leqslant k \leqslant k_{i1})\\ (i = 1, \ldots, I)\end{matrix} \tag{7-147}$$

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

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

Thus equation (7-144) becomes

$$d^2(f_kJ_k)/d\tau_k{}^2 = (1-\zeta_k)J_k - \xi_{ik}\left[\gamma_{ik}\sum_{l=k_{i0}}^{k_{i1}} w_l\Phi_{il}J_l + \varepsilon_{ik}(\partial B^0/\partial T)_k\sum_{k'=1}^{K} w_{k'}\psi_{k'}J_{k'}\right] - \xi_{ik}\varepsilon_{ik}[B_k{}^0 - (\partial B^0/\partial T)_kE_1] \tag{7-150}$$

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. espcially 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_ν, 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_ν and χ_ν, 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

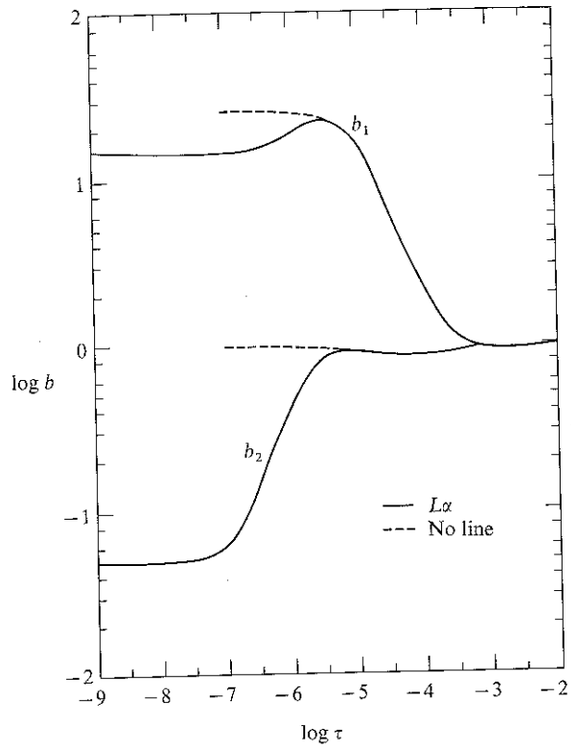

FIGURE 7-19
Non-LTE departure coefficients in the first two states of hydrogen for a model with $T_{\text{eff}} = 15{,}000°\text{K}$ and $\log g = 4$. *Dashed curves*: Lyman, Balmer and free–free continua only; *full curves*: Lα line included. The optical depth scale in the abscissa is measured just 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 Lα 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 Hα is taken into account. If we had computed the Hα 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

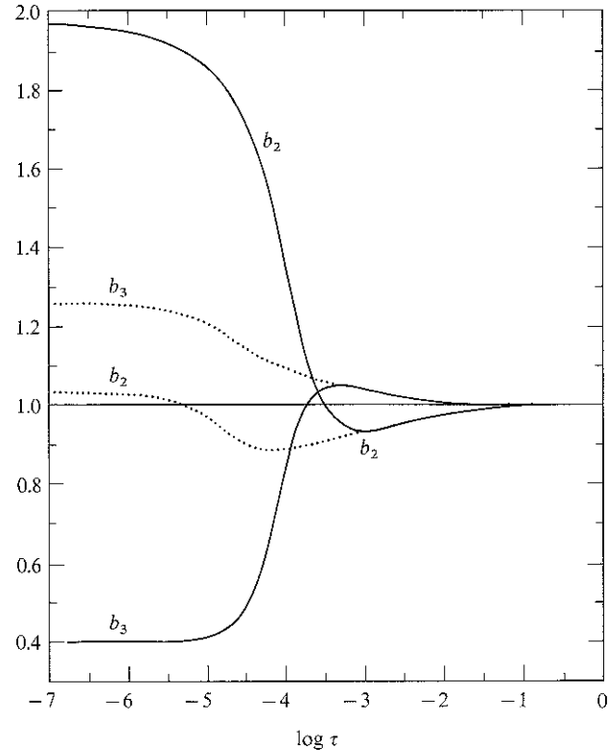

FIGURE 7-20
Non-LTE departure coefficients in the second and third level of hydrogen in a model with $T_{\text{eff}} = 15{,}000°\text{K}$ and $\log g = 4$. *Dotted curves*: Lyman, Balmer, Paschen, and free–free continua only; *full curves*: Hα included. Optical depth scale as in Figure 7-19. From (**41**),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

$$\boldsymbol{\psi}_d = (J_1, \ldots, J_K, N, T, n_e, n_1, \ldots, n_L)^T, \quad (d = 1, \ldots, D) \quad (7\text{-}151)$$

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 $\boldsymbol{\psi}_d$ as written; e.g., *given* all the J's, the n_i'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 $\boldsymbol{\psi}_d$ we require a total of $K + L + 3$ equations; we may choose these to be K transfer equations (J_k, $k = 1, \ldots, K$); 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 ($n_1, \ldots, n_L$), 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 $\boldsymbol{\psi}_d$ can be written in terms of the current (but imperfect) solution $\boldsymbol{\psi}_d{}^0$ as $\boldsymbol{\psi}_d = \boldsymbol{\psi}_d{}^0 + \boldsymbol{\delta\psi}_d$. We then choose $\boldsymbol{\delta\psi}_d$ so as to satisfy all constraints more closely; i.e., if $\mathbf{f}_d(\boldsymbol{\psi}_d) = 0$ represents the entire system of constraints, we demand that $\mathbf{f}_d(\boldsymbol{\psi}_d{}^0 + \boldsymbol{\delta\psi}_d) = 0$, and solve for $\boldsymbol{\delta\psi}_d$ by linearing the entire system:

$$\mathbf{f}_d(\boldsymbol{\psi}_d{}^0) + \sum_j \frac{\partial \mathbf{f}_d}{\partial \boldsymbol{\psi}_{d,j}} \cdot \boldsymbol{\delta\psi}_{d,j} = 0 \quad (7\text{-}152)$$

If we express the transfer equations as difference equations, and the constraints in terms of quadrature sums, the $\mathbf{f}$'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 $\{\nu_n\}$. We suppose that a starting solution $\boldsymbol{\psi}_d{}^0$, $(d = 1, \ldots, D)$, 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_ν and η_ν, and write

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

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_ν, 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

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

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:

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

which, when linearized, yields

$$-\delta N_d + \delta n_{e,d} + (1 + Y)\left[\delta n_{p,d} + \sum_{i=1}^{L_{\mathrm{II}}} \delta n_{i,d}(\mathrm{H})\right] = N_d - n_{e,d} - (1 + Y)\left[n_{p,d} + \sum_{i=1}^{L_{\mathrm{II}}} n_{i,d}(\mathrm{H})\right] \quad (7\text{-}156)$$

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

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

where

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

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 $\mathscr{A}\mathbf{n} = \mathscr{B}$.

The complete system is of the general form

$$-\mathbf{A}_d\,\delta\boldsymbol{\psi}_{d-1} + \mathbf{B}_d\,\delta\boldsymbol{\psi}_d - \mathbf{C}_d\,\delta\boldsymbol{\psi}_{d+1} = \mathbf{L}_d \quad (7\text{-}159)$$

which may be solved by the standard Feautrier elimination scheme.

Exercise 7-14: (a) Sketch the form of the **A**, **B**, and **C** matrices, indicating nonzero elements by x's as was done in §5-4 for the rate equations. Show that the **A** and **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 $\boldsymbol{\psi}_d{}^0$ are used; as $\mathbf{L}_d \to 0$, the corrections $\boldsymbol{\delta\psi}_d \to 0$. 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 $\boldsymbol{\delta\psi}$ 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 $\boldsymbol{\delta\psi}$'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°K $\leqslant T_{\text{eff}} \leqslant$ 30,000°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

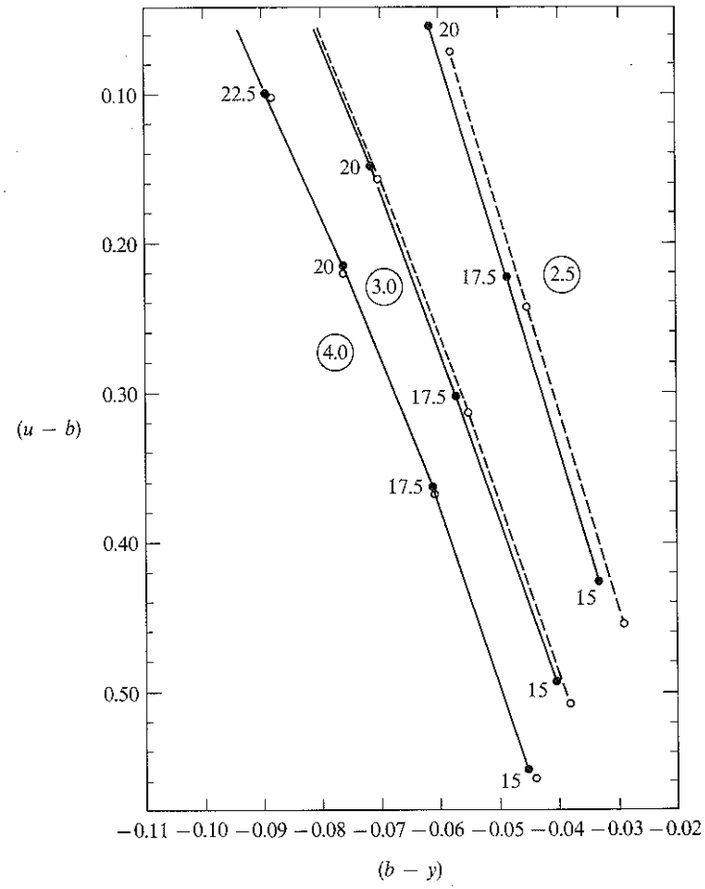

FIGURE 7-21
Theoretical Strömgren-system colors for LTE and non-LTE models. *Ordinate*: $(u - b)$; *abscissa*: $(b - y)$. LTE values are represented by open circles and dashed curves; non-LTE values by solid dots and curves. Curves are labeled with $\log g$, and individual models are labeled with $T_{\text{eff}}/10^3$.

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 $\lambda < 912$ Å 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

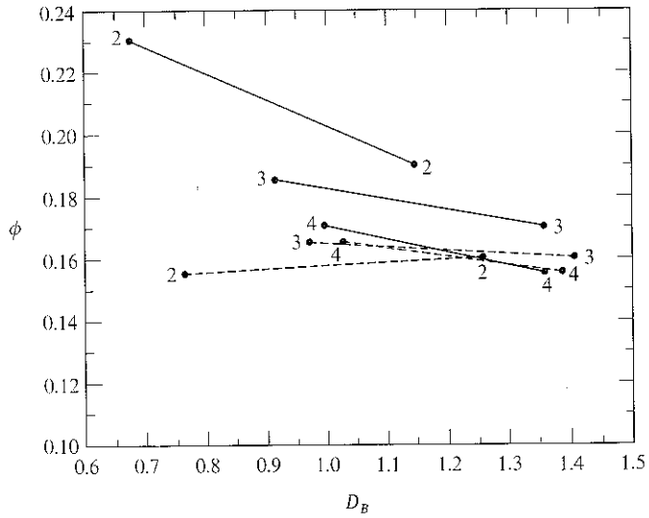

FIGURE 7-22
Theoretical continuum-jump parameters for LTE and non-LTE models of late B-type stars. *Ordinate*: $\phi \equiv D_P/D_B$; *abscissa*: Balmer jump D_B. *Solid dots and curves*: non-LTE values; *open dots and dashed curves*: LTE values. Curves are labeled with log g.

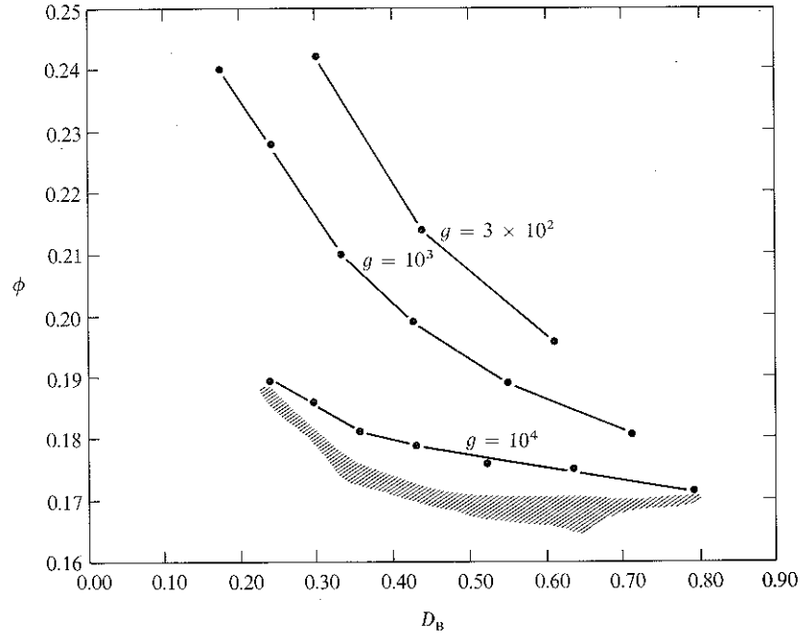

FIGURE 7-23
Same as Figure 7-22 for middle B-type stars. Shaded area contains predictions of LTE models at various values of g; curves (labeled with surface-gravity g) show non-LTE values.

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 ($T_{\text{eff}} \gtrsim 30{,}000$°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 λ227 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

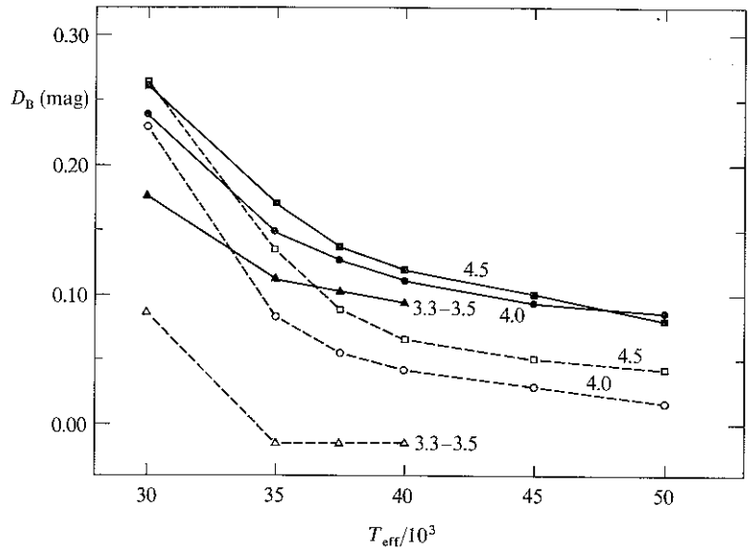

FIGURE 7-24
Theoretical Balmer jumps for O-stars. *Ordinate*: D_B in magnitudes; *abscissa*: $T_{eff}/10^3$. *Solid symbols and curves*: non-LTE models; *open symbols and dashed curves*: LTE models. Curves are labeled with log g. From (**45**), by permission.

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 CI 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 $T_0/T_{eff} = 0.811$. 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 opague, 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

$$n_1^* \int_{\nu_0}^{\infty} \alpha_\nu B_\nu(T_0)\, d\nu = n_1^* \int_{\nu_0}^{\infty} \alpha_\nu J_\nu\, d\nu = n_1^* W \int_{\nu_0}^{\infty} \alpha_\nu B_\nu(T_R)\, d\nu \quad (7\text{-}160)$$

where the last equality introduces a parametric representation of J_ν 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 $T_R = T_{eff}$, 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

$$n_1^* \int_{\nu_0}^{\infty} \alpha_\nu B_\nu(T_0)\, d\nu = b_1 n_1^* \int_{\nu_0}^{\infty} \alpha_\nu J_\nu\, d\nu = b_1 n_1^* W \int_{\nu_0}^{\infty} \alpha_\nu B_\nu(T_R)\, d\nu \quad (7\text{-}161)$$

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

$$b_1 = \int_{\nu_0}^{\infty} (\alpha_\nu B_\nu/h\nu)\, d\nu \Big/ \int_{\nu_0}^{\infty} (\alpha_\nu J_\nu/h\nu)\, d\nu \quad (7\text{-}162)$$

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

$$\left[\int_{\nu_0}^{\infty} \alpha_\nu B_\nu(T_0)\, d\nu \Big/ \int_{\nu_0}^{\infty} \alpha_\nu B_\nu(T_0)\nu^{-1}\, d\nu\right] = \left[\int_{\nu_0}^{\infty} \alpha_\nu B_\nu(T_R)\, d\nu \Big/ \int_{\nu_0}^{\infty} \alpha_\nu B_\nu(T_R)\nu^{-1}\, d\nu\right] \quad (7\text{-}163)$$

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 $\bar{\tau} \approx 1$, to a value $T_{\min}$, approximately equal to the grey-body boundary-temperature, and then rise to a value T_0, with $T_{\min} < T_0 \lesssim T_{\text{eff}}$. 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, $T_{\text{eff}} = 5900°\text{K}$, and Cayrel estimated $T_0 = T_R \approx 5600°\text{K}$ compared to $T_{\min} \approx 4800°\text{K}$ (grey value). From a detailed calculation of a solar radiative equilibrium model, allowing for departures from LTE in the H^- ion, Feautrier obtained (**211**) $T_{\min} \approx 4700°\text{K}$ and $T_0 \approx 5200°\text{K}$. 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 $T_{\text{eff}} = 15{,}000°\text{K}$, $\log g = 4$, had a plateau at 10,400°K, where the Lyman continuum was optically thick and all others were transparent, followed by a drop to $T_0 = 9400°\text{K}$ 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°K and then a rise outward to $T_0 \approx 10{,}350°\text{K}$, 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

$$S_i = (1 - \varepsilon_i)\bar{J}_i + \varepsilon_i B_\nu \quad (7\text{-}164)$$

where $\bar{J}$ denotes the average of J_ν against the line-profile ϕ_ν. 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 ($\varepsilon = 1$) there is a *large* drop in boundary temperature, but when $\varepsilon \ll 1$, 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 $T_{\min} \approx 4330°\text{K}$, 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 $T_0 \approx 10{,}350°\text{K}$ (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 ε is $\ll 1$. 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

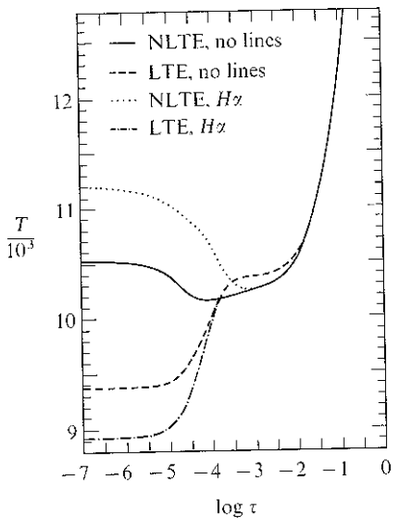

FIGURE 7-25
Temperature distribution for LTE and non-LTE models with $T_{\text{eff}} = 15{,}000°\text{K}$ and $\log g = 4$. The atmosphere is composed of hydrogen, which is represented by a schematic model atom with three bound levels and continuum. This model atom accounts for the Lyman, Balmer, Paschen, and free–free continua, and the Hα 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°K (the continuum-only value) to 11,200°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 outweighted 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 $T_{\text{eff}} = 30{,}000°\text{K}$, $\log g = 4$, (**436**) including the $L\alpha$, $L\beta$, $L\gamma$, $H\alpha$, $H\beta$, 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 problem of non-LTE line-blanketing in stellar atmospheres has barely been approached, and a large amount of work remains to be done.

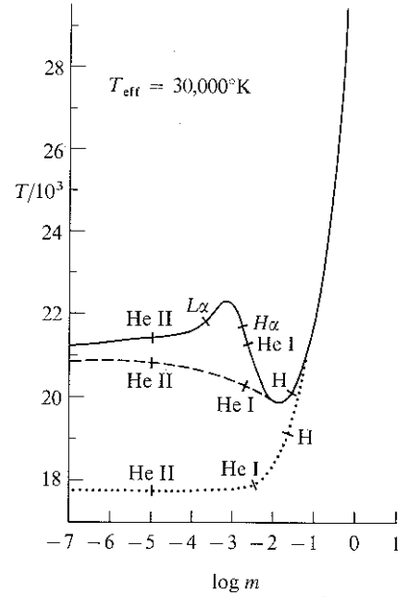

FIGURE 7-26
Temperature distributions in models with $T_{\text{eff}} = 30{,}000°\text{K}$. *Ordinate*: $T/10^3$; *abscissa*: logarithm of column mass in gm cm^{-2}. *Solid curve*: non-LTE model including lines; *dashed curve*: non-LTE continuum-only model; *dotted curve*: LTE model. Cross-bars on curves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

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

$$\mu(\partial I_\nu/\partial r) + r^{-1}(1 - \mu^2)(\partial I_\nu/\partial \mu) = \eta_\nu - \chi_\nu I_\nu \tag{7-165}$$

with moments (cf. §2-4)

$$r^{-2}[\partial(r^2 H_\nu)/\partial r] = \eta_\nu - \chi_\nu J_\nu \tag{7-166}$$

and

$$(\partial K_\nu/\partial r) + r^{-1}(3K_\nu - J_\nu) = -\chi_\nu H_\nu \tag{7-167a}$$

or

$$[\partial(f_\nu J_\nu)/\partial r] + r^{-1}(3f_\nu - 1)J_\nu = -\chi_\nu H_\nu \tag{7-167b}$$

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_ν 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_ν and J_ν accurately, as is possible in the planar case. But now, even though $f_\nu \to \frac{1}{3}$ at great depth where the radiation field is homogeneous, $f_\nu \to 1$ 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 $\mu_* \leqslant \mu \leqslant 1$ [where $\mu_* \equiv (1 - r_*^2/r^2)^{\frac{1}{2}}$] separately from those computed on the range $-1 \leqslant \mu \leqslant \mu_*$ [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 $\chi_\nu \equiv \chi$, and integrate equations (7-166) and (7-167) over all frequency, omitting the subscript ν to denote integrated quantities, we obtain

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

and

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

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

$$\int_0^\infty \eta_\nu \, d\nu = \int_0^\infty \chi_\nu J_\nu \, d\nu = \chi J \tag{7-170}$$

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

$$r^2 H = H_0 = L/16\pi^2 \tag{7-171}$$

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

$$\tau(r) = \int_r^R \chi(r') \, dr' \tag{7-172}$$

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

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

which yields the integral

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

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

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

a result obtained by Chandrasekhar (**148**). If, further, we impose the assumption of LTE, then $\eta_\nu = \chi_\nu B_\nu = \chi B_\nu$ 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 $f \to 1$, in which case equation (7-169) becomes

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

which yields

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

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

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

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

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 $\rho v r^2 =$ 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

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

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

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

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

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

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

$$T(\tau) = T_1 \tau^{\frac{1}{2(n-1)}} \{[\tau + (n+1)/3(n-1)]/[1 + (n+1)/3(n-1)]\}^{\frac{1}{4}} \tag{7-182}$$

Equation (7-182) reveals an important characteristic difference between the planar and spherical cases, namely $T \to 0$ as $\tau \to 0$ 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: $z = r\cos\theta, \quad p = r\sin\theta$, 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

$$(\partial I_\nu/\partial z) = \eta_\nu - \chi_\nu I_\nu \tag{7-183}$$

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

$$I_\nu(p, \infty) = \int_{-\infty}^{\infty} B_\nu[T(p, z)] \exp[-\tau(p, z)]\chi(p, z)\, dz \tag{7-184}$$

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

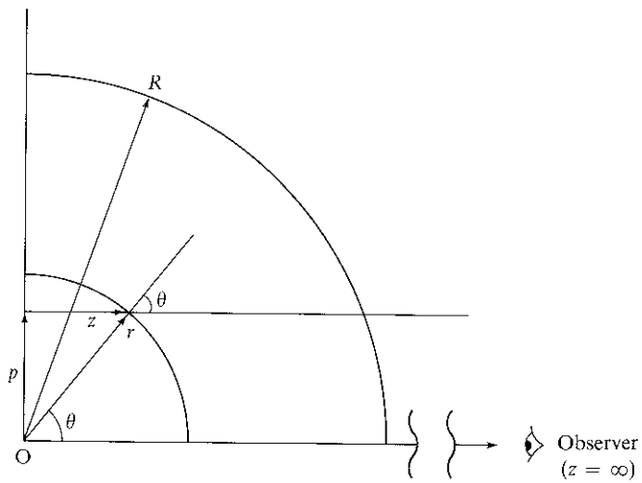

FIGURE 7-27
Coordinate systems for solution of transfer equation in spherical geometry.

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

$$\mathscr{F}_\nu = 2\pi D^{-2} \int_0^\infty I_\nu(p, \infty) p \, dp \tag{7-185}$$

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

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

Then

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

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

$$\mathscr{F}_\nu = \pi (R_1/D)^2 \int_0^\infty B_\nu[T(\tau)] \tau^{-\left(\frac{2}{n-1}\right)} \Phi_n(\tau) \, d\tau \tag{7-188}$$

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

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

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 ($n \to \infty$ 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 $\lambda_c = 5000$ Å for various models with $T_1 = 50{,}000°$K. He finds that, at $n = \infty$, $T_c/10^4 = 5$, and at $n =$ (10, 5, 3, and 2), $T_c/10^4 =$ (4.3, 3.5, 2.2, and 1.2) respectively, which shows that increasing atmospheric extent produces effects that simulate lower atmospheric temperatures. Indeed, a model with $T_1 = 50{,}000°$K and $n = 3$ has a flux distribution nearly identical to

one with $T_1 = 30{,}000°$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 $\lambda_c = 5000$ Å, 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

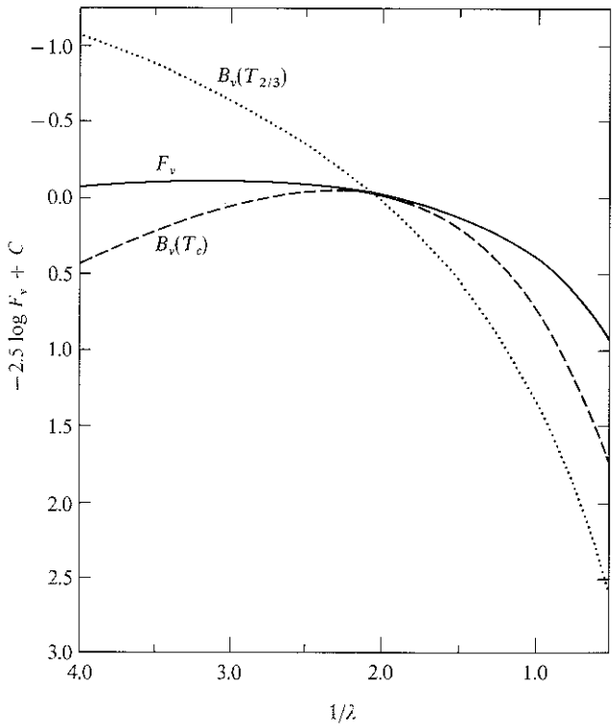

FIGURE 7-28
Solid curve: emergent flux F_ν from spherical grey atmosphere with $T_1 = 5 \times 10^4$ °K and $n = 2$. *Dashed curve*: black-body curve at color temperature T_c that matches slope of F_ν at $\lambda 5000$ Å; note ultraviolet and infrared excesses of F_ν relative to $B_\nu(T_c)$. *Dotted 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 ($f_K \to 1$) 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 ν, and introduce the radial optical depth scale $d\tau_\nu = -\chi_\nu\, dr$. Then the moment equations are

$$\partial(r^2 H_\nu)/\partial\tau_\nu = r^2(J_\nu - S_\nu) \tag{7-190}$$

and

$$[\partial(f_\nu J_\nu)/\partial\tau_\nu] - (3f_\nu - 1)J_\nu/(\chi_\nu r) = H_\nu \tag{7-191}$$

where S_ν is assumed to have a general form

$$S_\nu = \alpha_\nu \int R(r; \nu', \nu) J_{\nu'}\, d\nu' + \beta_\nu \tag{7-192}$$

For a pure continuum problem, we would have only a coherent scattering term and S_ν 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 $(\chi_\nu r)^{-1}$ tends to diverge strongly at the surface (recall χ_ν 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_ν, defined by

$$\ln(r^2 q_\nu) = \int_{r_c}^{r} [(3f_\nu - 1)/(r' f_\nu)]\, dr' + \ln r_c^{\,2} \tag{7-193}$$

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

$$\partial(f_\nu q_\nu r^2 J_\nu)/\partial\tau_\nu = q_\nu r^2 H_\nu \tag{7-194}$$

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

$$\frac{\partial}{\partial\tau_\nu}\left[\frac{1}{q_\nu}\frac{\partial(f_\nu q_\nu r^2 J_\nu)}{\partial\tau_\nu}\right] = r^2(J_\nu - S_\nu) \tag{7-195}$$

or, introducing a new variable $dX_\nu = -q_\nu\chi_\nu\, dr = q_\nu\, d\tau_\nu$,

$$\partial^2(f_\nu q_\nu r^2 J_\nu)/\partial X_\nu^{\,2} = q_\nu^{-1} r^2(J_\nu - S_\nu) \tag{7-196}$$

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

$$h_\nu \equiv \int_0^1 I(R, \mu, \nu)\mu\, d\mu \Big/ \int_0^1 I(R, \mu, \nu)\, d\mu \tag{7-197}$$

so that from (7-194)

$$[\partial(f_\nu q_\nu r^2 J_\nu)/\partial X_\nu]_{r=R} = h_\nu (r^2 J_\nu)_{r=R} \tag{7-198}$$

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

$$H_\nu(r_c) = \frac{1}{3}(\chi_\nu^{-1}|\partial B_\nu/\partial r|)_{r_c} \tag{7-199}$$

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

$$[\partial(f_\nu q_\nu r^2 J_\nu)/\partial X_\nu]_{r=r_c} = r_c^2 H_c \left[\chi_\nu^{-1}(\partial B_\nu/\partial T)\Big/\int_0^\infty \chi_\nu^{-1}(\partial B_\nu/\partial T)\,d\nu\right]_{r=r_c} \tag{7-200}$$

The diffusion approximation, and hence equations (7-199) and (7-200), will be valid when the photon mean free path $\chi_\nu^{-1} \ll \varepsilon R$ 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 $\{r_d\}$, $(d = 1, \ldots, D)$, where $R = r_1 > r_2 > \cdots > r_D = r_c$, and a frequency mesh $\{\nu_n\}$, $(n = 1, \ldots, N)$, 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_ν; 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, $(i = 1, \ldots, I)$, 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 $r_d \geqslant p_i$, and these intersections define a mesh of z-points $\{z_{di}\}$, $(d = 1, \ldots, D_i)$. Here $D_i = D + C + 1 - i$ for $i > C$, $D_i = D$ for $i \leqslant C$, 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

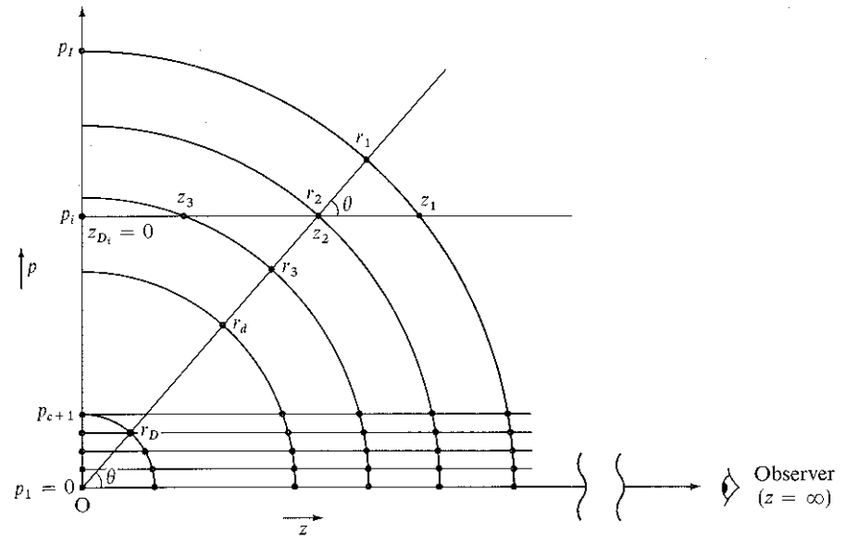

FIGURE 7-29
Discrete (p, z) mesh used in solution of spherical transfer equation. The impact parameters $\{p_i\}$ are chosen to be parallel to the central ray, and tangent to spherical shells chosen to describe the depth-variation of physical properties of the envelope. The intersections of the rays 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 $I_\nu(z_{di}, p_i)$ $(i = 1, \ldots, I_d)$ is *equivalent to knowledge of the* μ*-variation of* $I_\nu(r_d, \mu)$ *on the mesh* $\{\mu_{di}\}$, $(i = 1, \ldots, I_d)$, which spans the interval $1 \geqslant \mu \geqslant 0$. 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

$$\pm[\partial I^\pm(z, p_i, \nu)/\partial z] = \eta(r, \nu) - \chi(r, \nu) I^\pm(z, p_i, \nu) \tag{7-201}$$

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 $d\tau(z, p_i, \nu) \equiv -\chi(r, \nu)\,dz$, setting $S(r, \nu) \equiv \eta(r, \nu)/\chi(r, \nu)$ (assumed to be *known* values), and introducing the

mean-intensity-like and flux-like variables

$$u(z, p_i, \nu) \equiv \frac{1}{2}\left[I^+(z, p_i, \nu) + I^-(z, p_i, \nu)\right] \tag{7-202}$$

and

$$v(z, p_i, \nu) \equiv \frac{1}{2}\left[I^+(z, p_i, \nu) - I^-(z, p_i, \nu)\right] \tag{7-203}$$

we obtain the second-order system

$$\left[\partial^2 u(z, p_i, \nu)/\partial\tau^2(z, p_i, \nu)\right] = u(z, p_i, \nu) - S[r(z, p_i), \nu] \tag{7-204}$$

with an upper boundary condition

$$\partial u(z, p_i, \nu)/\partial\tau(z, p_i, \nu)\big|_{z_{\max}} = u(z_{\max}, p_i, \nu) \tag{7-205}$$

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 $v(0, p_i, \nu) \equiv 0$, hence

$$\partial u(z, p_i, \nu)/\partial\tau(z, p_i, \nu)\big|_{z=0} = 0 \tag{7-206}$$

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 $u_{din} \equiv u(z_d, p_i, \nu_n)$ we construct the moments, as described above,

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

and

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

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 extending the difference equation for equation (7-204) beyond the central plane $z_{Di} = 0$, 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

$$\mathbf{T}_{in}\mathbf{u}_{in} = \mathbf{U}_{in}\bar{\mathbf{J}} + \mathbf{W}_{in} \qquad \begin{matrix}(i = 1, \ldots, I)\\ (n = 1, \ldots, N)\end{matrix} \tag{7-209}$$

where $\bar{\mathbf{J}}$ describes the depth-variation of $\bar{J}$ [cf. equation (6-46)] on the range $(d = 1, \ldots, D_i)$, and $\mathbf{u}_{in}$ the variation of $u(z, p_i, \nu_n)$ along the ray. This system can be solved for $\mathbf{u}_{in} = \mathbf{C}_{in}\bar{\mathbf{J}} + \mathbf{D}_{in}$, and this solution can be substituted into the equation defining $\bar{\mathbf{J}}$ [cf. equation (7-207)] to yield a final system for $\bar{\mathbf{J}}$; 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 $\mathscr{M}$ is the mass of the star, the equation of hydrostatic equilibrium becomes

$$(dp_g/dr) = -(\rho G\mathscr{M}/r^2) + (4\pi/c)\int_0^\infty \chi_\nu H_\nu\, d\nu \tag{7-210}$$

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

$$\gamma \equiv (\chi_R H)^{-1}\int_0^\infty \chi_\nu H_\nu\, d\nu \tag{7-211}$$

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

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 $\Gamma = \gamma\chi_R L/(4\pi cG\mathscr{M}\rho)$ 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 $(L/\mathscr{M})$ at which radiation forces just balance gravity forces, namely $(L/\mathscr{M})_{\text{crit}} = 3.8 \times 10^4\ (L_\odot/\mathscr{M}_\odot)$. (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 $r_{\frac{2}{3}}$, at which $\tau_R = \frac{2}{3}$, as a characteristic radius.

The effects of radiation forces are quite important in the planetary nebula stars, as estimates of L and $\mathscr{M}$ 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 $g_{\text{eff}}^{-1} = [(1 - \Gamma)G\mathscr{M}/R^2]^{-1}$, 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 $\lambda_c = 5000$ Å, 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, $\mathscr{M}$, 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 $\Gamma_{\max} \approx 0.995$). 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

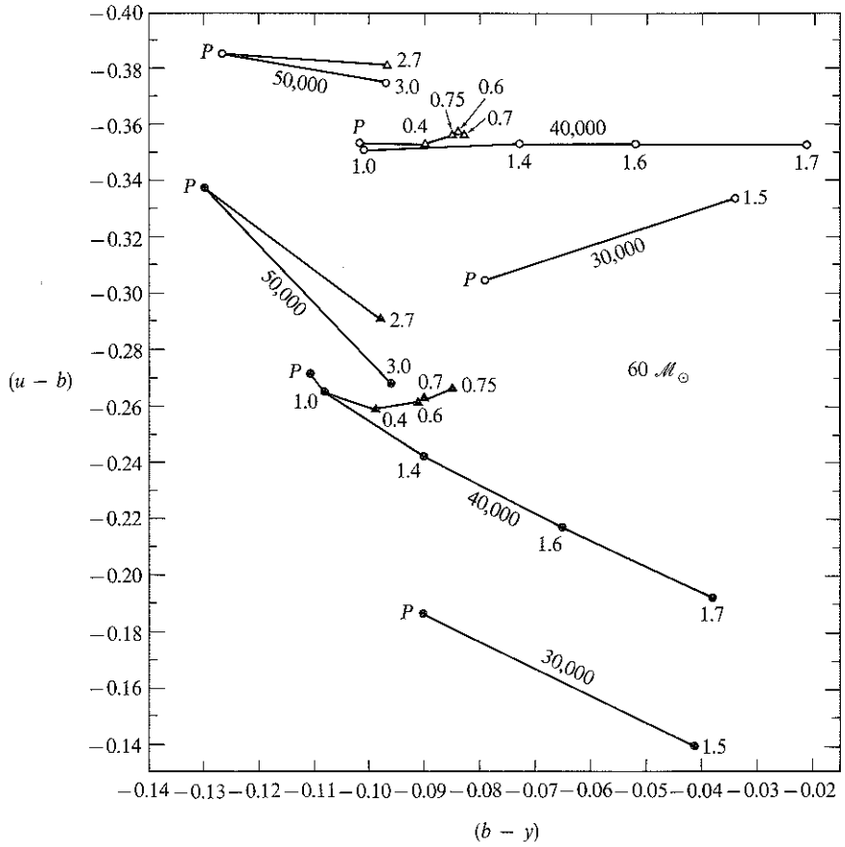

FIGURE 7-30
Strömgren-system $(u - b)$ and $(b - y)$ colors for models with $\mathscr{M} = 60\ \mathscr{M}_\odot$ *Open symbols*: LTE models; *filled symbols*: non-LTE models. Models with equal values of T_{eff} are joined on curves. Curves are labeled with parameters describing radiation-force multiplier γ. *Circles*: $\gamma \equiv \gamma_1$, 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

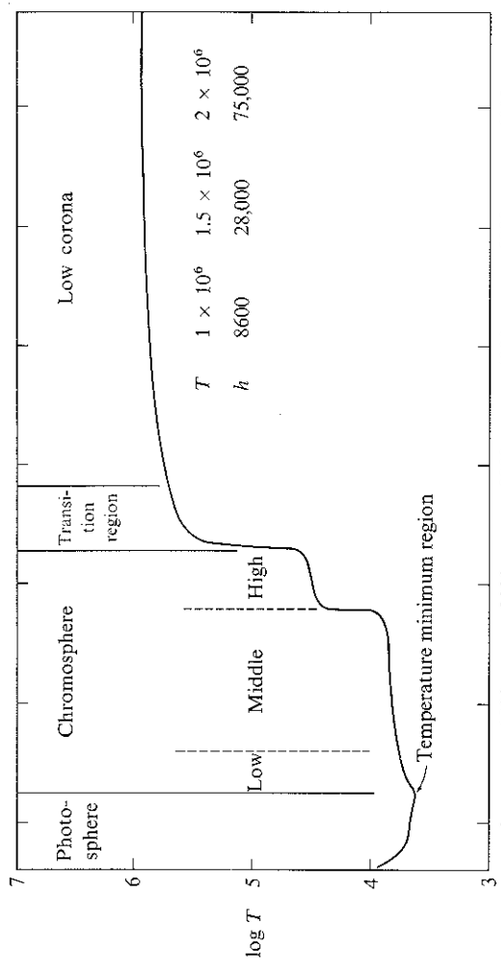

FIGURE 7-31
Temperature as a function of height in the solar atmosphere. The zero-point of the height-scale in this plot refers to unit *tangential* optical depth at the limb (radial optical depth $\tau_{5000} \approx 3.7 \times 10^{-3}$), and lies at approximately +340 km on the scale of Figure 7-32 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 Å $\lesssim \lambda \lesssim 350\ \mu$ 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 $\lambda \gtrsim 1900$ Å are dominantly in absorption, and a transition to emission lines occurs in the range 1700 Å $\lesssim \lambda \leqslant 1900$ Å. The continuum on the range 1525 Å $\lesssim \lambda \lesssim 1685$ Å comes from the photosphere–chromosphere transition region, and for $\lambda <$ 1525 Å is dominantly chromospheric. On the range 504 Å $\leqslant \lambda \lesssim 912$ Å the Lyman continuum is in emission; for $\lambda \lesssim 504$ Å the emission is from the continuum of He I. Chromospheric emission lines extend at least down to 288 Å, through the resonance series of He II. Coronal emission lines begin to appear prominently at about 800 Å, and extend down into the X-ray region. In the infrared, one sees most deeply into the atmosphere near 1.6 μ, the minimum in the H^- opacity. At longer wavelengths, the temperature minimum region is encountered near 300 μ; 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 $I_\nu(0, \mu)$ comes from $S_\nu(\tau_\nu \approx \mu)$; hence by scanning from center to limb,

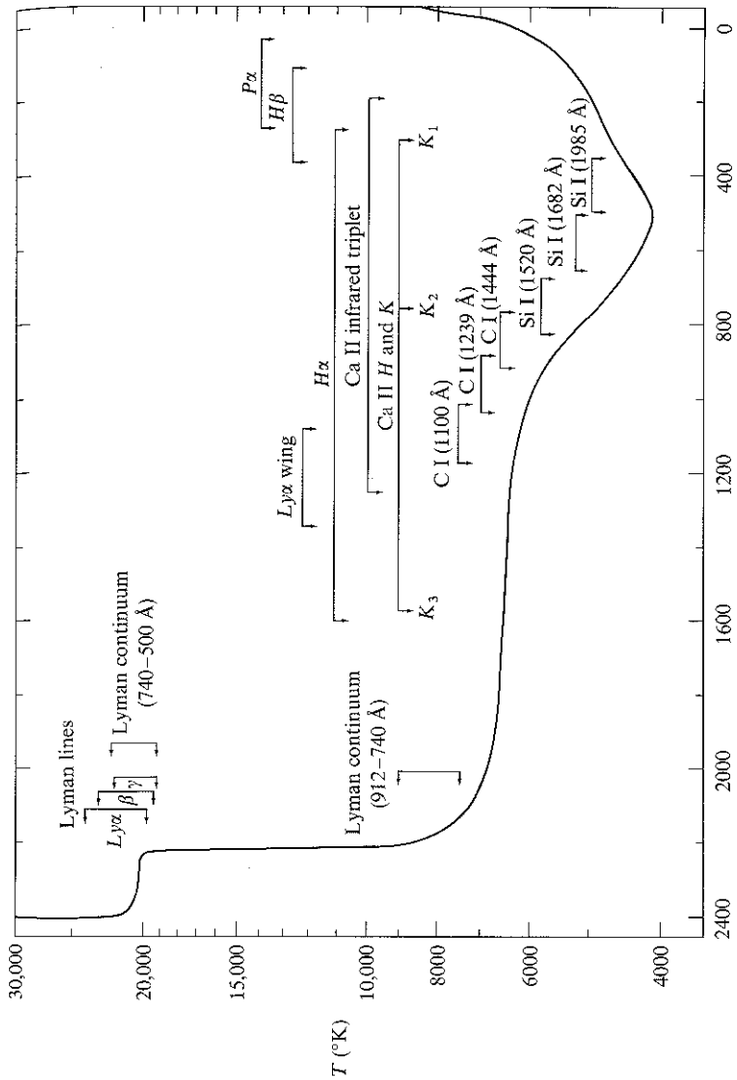

FIGURE 7-32
Temperature as a function of height in the solar atmosphere. Unit optical depth for various lines and continua (over their entire profiles) are indicated by lines with arrows. From **(645)**, by permission.

$0 \leqslant \mu \leqslant 1$, we sample depths $0 \lesssim \tau_\nu \lesssim 1$. Suppose we assume that the source function is the Planck function, and that we can write $B_\nu(\tau_\nu)/I_\nu(0, 1) \equiv b_\nu(\tau_\nu)$. Then

$$\Phi_\nu(\mu) = [I_\nu(0, \mu)/I_\nu(0, 1)] = \int_0^\infty b_\nu(\tau_\nu) \exp(-\tau_\nu/\mu)\, d\tau_\nu/\mu \quad (7\text{-}213)$$

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

$$b_\nu = \sum_k a_k \tau_\nu{}^k \quad (7\text{-}214)$$

then we find

$$\Phi_\nu(\mu) = \sum_k a_k k!\, \mu^k \quad (7\text{-}215)$$

Thus a fit of equation (7-215) to the data yields the a_k's, hence b_ν and, ultimately, if we know $I_\nu(0, 1)$ we obtain $B_\nu(\tau_\nu)$, and hence $T(\tau_\nu)$. 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 $\tau_\nu > 1$, 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 $0.05 \lesssim \tau_\nu \lesssim 1$. 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 $\Phi_\nu(\mu)$ but show wild oscillations in $b_\nu(\tau_\nu)$. 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 $\Phi_\nu(\mu)$ equally well; in general these yield different values of $b_\nu(\tau_\nu)$ and hence of $T(\tau_\nu)$. Among three-term solutions, discrepancies in the temperature of the order of $\pm 200°$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 $\pm 200°$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 $T(\tau_\nu)$ from one frequency to another, which can be done if we know the frequency variation of the absorption coefficient—i.e., $\chi(\nu)/\chi(\nu_{\text{std}})$. Alternatively, we can demand that the different $T(\tau_\nu)$ 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 semiempirical 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(km)	T(°K)	p_g(dynes cm^{-2})	p_e(dynes cm^{-2})	$n_p/[n_p + n(\mathrm{H})]$
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 (λ6374) and [Fe XIV] (λ5303). Modern determinations of the temperature in the corona are in agreement with one another and give a typical temperature of about 1.5×10^6 °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 $\gg T_{\mathrm{eff}}$!). 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^s 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 $\pm 100°$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°K to 30,000°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} $\mathscr{M}_\odot$/year) is well established. These winds make the solar wind (mass-loss rate $\approx 10^{-14}$ $\mathscr{M}_\odot$/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 inacessible 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_ν (the flux in the line) relative to F_c (the flux in the continuum) and describe the profile in terms of its *absorption depth*

$$A_\nu \equiv 1 - (F_\nu/F_c) \tag{8-1}$$

or *residual flux*

$$R_\nu \equiv (F_\nu/F_c) = 1 - A_\nu \tag{8-2}$$

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 $I_\nu(0, \mu)$, the emergent specific intensity, in units of the nearby continuum intensity $I_c(0, \mu)$, and write

$$a_\nu(\mu) \equiv 1 - [I_\nu(0, \mu)/I_c(0, \mu)] \tag{8-3}$$

or
$$r_\nu(\mu) \equiv I_\nu(0, \mu)/I_c(0, \mu) = 1 - a_\nu(\mu) \tag{8-4}$$

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

$$W_\nu \equiv \int_0^\infty A_\nu \, d\nu \tag{8-5}$$

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

$$W_\lambda \equiv \int_0^\infty A_\nu \, d\lambda \tag{8-6}$$

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

$$W_\nu(\mu) \equiv \int_0^\infty a_\nu(\mu) \, d\nu \tag{8-7}$$

or
$$W_\lambda(\mu) \equiv \int_0^\infty a_\nu(\mu) \, d\lambda \tag{8-8}$$

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 ν, and suppose that we know the continuous absorption and scattering coefficients, κ_ν and σ_ν, and the line absorption coefficient $\chi_l\phi_\nu$ as a function of depth. We could then construct the line optical depth scale

$$\tau_\nu \equiv \int_z^{z_{\max}} (\kappa_\nu + \sigma_\nu + \chi_l\phi_\nu) \, dz' \tag{8-9}$$

and continuum optical depth scale

$$\tau_c \equiv \int_z^{z_{\max}} (\kappa_\nu + \sigma_\nu) \, dz' \tag{8-10}$$

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

$$F_\nu = 2 \int_0^\infty S_\nu(\tau_\nu) E_2(\tau_\nu) \, d\tau_\nu \tag{8-11}$$

$$F_c = 2 \int_0^\infty S_c(\tau_c) E_2(\tau_c) \, d\tau_c \tag{8-12}$$

and hence A_ν 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_ν known beforehand (namely $S_\nu = B_\nu$). 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 $R(\nu', \mathbf{n}'; \nu, \mathbf{n})$. Others may be destroyed by collisional deexcitations 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

$$\mu(\partial I_\nu/\partial z) = -(\kappa_\nu + \sigma_\nu + \chi_l\phi_\nu)I_\nu + \kappa_\nu B_\nu + \sigma_\nu J_\nu + \bar{\gamma}\chi_l \oint (d\omega/4\pi) \int_0^\infty d\nu' I(\nu', \mathbf{n}')R(\nu', \mathbf{n}'; \nu, \mathbf{n}) + \bar{\varepsilon}\chi_l\phi_\nu B_\nu \quad (8\text{-}13)$$

Here the coupling coefficients $\bar{\gamma}$ and $\bar{\varepsilon}$ 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_ν. This will be treated in Chapter 9.

(b) We must be able to specify the coupling between photons and the material via the parameters $\bar{\gamma}$ and $\bar{\varepsilon}$. 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 $\bar{\gamma}$ and $\bar{\varepsilon}$ 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 $R(\nu', \mathbf{n}'; \nu, \mathbf{n})$, 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 $\bar{\gamma}$ and $\bar{\varepsilon}$, 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, 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.

9

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

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

which satisfies the fundamental *reciprocity relation*

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

The quantity

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

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

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

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

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

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*

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

from which the power spectrum is obtained by the relation

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

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

$$\ddot{x} = -\omega_0^2 x - \gamma \dot{x} \tag{9-8}$$

where γ is the *classical damping constant*

$$\gamma \equiv (2e^2\omega_0^2/3mc^3) \tag{9-9}$$

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

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

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

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

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

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

The energy spectrum of the oscillator is then

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

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

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

is

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

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

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

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

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

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

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

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

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

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

$$\Gamma_U = \sum_{i<U} A_{Ui} \tag{9-20}$$

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

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

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 $x_0 \equiv \omega - \omega_0$,

$$x' = x - x_0 \tag{9-22}$$

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

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

This integral may be evaluated by contour integration using the *residue theorem*, taking into account the poles at $z = \pm i\delta_U$ 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

$$\Gamma \equiv \Gamma_L + \Gamma_U \tag{9-24}$$

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

$$\int_{-\infty}^{\infty} \alpha_\nu\, d\nu = (\pi e^2/mc) f \tag{9-25}$$

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

$$\alpha_\nu = \left(\frac{\pi e^2}{mc}\right) f \frac{(\Gamma/4\pi^2)}{(\nu - \nu_0)^2 + (\Gamma/4\pi)^2} \tag{9-26}$$

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

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

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

$$\alpha_\nu = \int_{-\infty}^{\infty} \alpha(\nu - \xi\nu/c) W(\xi)\, d\xi \tag{9-28}$$

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

$$v \equiv (\nu - \nu_0)/\Delta\nu_D \tag{9-29}$$

$$y \equiv (\Delta\nu/\Delta\nu_D) = (\xi/\xi_0) \tag{9-30}$$

and

$$a \equiv (\Gamma/4\pi\, \Delta\nu_D) \tag{9-31}$$

where $\Delta\nu_D$ is the *Doppler width* of the line

$$\Delta\nu_D \equiv (\xi_0\nu_0/c) \tag{9-32}$$

we find the absorption coefficient can be written as

$$\alpha_\nu = (\pi^{\frac{1}{2}} e^2 f/mc\, \Delta\nu_D) H(a, v) \tag{9-33}$$

where

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

is known as the *Voigt function*. In deriving equation (9-34), the approximation that $\xi\nu/c \backsimeq \xi\nu_0/c$ (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

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

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

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

But the cosine transformation of the Gaussian is

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

from which we see that

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

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

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

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

From equation (9-37) it follows immediately that

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

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

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

and using the sine transform of the Gaussian

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

we have

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

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 $n \leqslant 4$.

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 $H_n(v) = -[d^2 H_{n-2}(v)/dv^2]/[n(n-1)]$ 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

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

where the first term applies in the line core, $v \leqslant v^*$, and the second in the line wing, $v \geqslant v^*$, 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

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

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

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

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

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

Computation of the integral, with normalization, yields

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

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

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

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

$$\tau^{-1} = \pi\rho_0^2 Nv \tag{9-51}$$

and

$$\Gamma = 2\pi\rho_0^2 Nv \tag{9-52}$$

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

$$\Delta\omega = C_p/r^p \tag{9-53}$$

where $r(t) = (\rho^2 + v^2t^2)^{\frac{1}{2}}$, $t = 0$ 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

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

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

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

where

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

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

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

and the corresponding value for the damping constant is

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

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

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

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),

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

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

$$\eta(t, s) \equiv \eta(t + s) - \eta(t) \tag{9-61}$$

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

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

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

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.,

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

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

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

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

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

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

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

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

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

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

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

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

Thus Lindholm theory yields a Lorentz profile with a damping *width*

$$\Gamma = 2Nv\sigma_R \tag{9-71}$$

and a *line shift*
$$\Delta\omega_0 = Nv\sigma_I \tag{9-72}$$

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

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

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	$C_4^{\frac{3}{5}} v^{\frac{1}{5}} N$	8.08	$C_6^{\frac{3}{5}} v^{\frac{3}{5}} N$
$\Delta\omega_0$		9.85		2.94	
$\Gamma/\Delta\omega_0$		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

$$C_3 = (e^2 f_{1n}/2m\omega_{1n}) \tag{9-74}$$

[see (**662**; **112**, 231)]. A quantum mechanical calculation gives a Γ slightly different from the Lindholm value, namely

$$\Gamma_3 = (16\pi n_1 C_3/3) = (4n_1 e^2 f_{1n}/3mv_{1n}) \tag{9-75}$$

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)]

$$\Delta\omega = C_6/r^6 = e^2\alpha a_0^2[\bar{R}_u^2 - \bar{R}_l^2]/(\hbar r^6) \tag{9-76}$$

where α is the polarizability of hydrogen and $\bar{R}^2$ is the mean square radius of the two levels. Quantum mechanical results are sometimes available for $\bar{R}^2$; 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

$$\tau_s = (\psi_p\rho_0/v) \tag{9-77}$$

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 $(\tau_s/\tau) = \frac{3}{4}\psi_p(\rho_0/r_0)^3$. 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 $\rho \to \infty$, 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.,

$$\Delta\omega_W = (v^p/C_p\psi_p{}^p)^{1/(p-1)} \tag{9-78}$$

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.,

$$\omega_s = 1/\tau_s \ll \omega_{ij} = |E_i - E_j|/\hbar \tag{9-79}$$

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 motions 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

$$I(\Delta\omega)\,d(\Delta\omega) \propto W(r)[dr/d(\Delta\omega)]\,d(\Delta\omega) \tag{9-80}$$

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

$$W(r) = \left[1 - \int_0^r W(x)\,dx\right](4\pi r^2 N)\,dr \tag{9-81}$$

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

$$\frac{d}{dr}\left[\frac{W(r)}{4\pi r^2 N}\right] = -(4\pi r^2 N)\left[\frac{W(r)}{4\pi r^2 N}\right] \tag{9-82}$$

and thus by integration and normalization,

$$W(r) = 4\pi r^2 N \exp\left(-\frac{4}{3}\pi r^3 N\right) \tag{9-83}$$

It is customary to adopt the *mean interparticle distance* $r_0 = (\frac{3}{4}\pi N)^{\frac{1}{3}}$ as the *reference distance* at which a perturber produces the *normal frequency shift* $\Delta\omega_0 = C_p/r_0^p$. Then

$$(\Delta\omega/\Delta\omega_0) = (r_0/r)^p \tag{9-84}$$

and equation (9-83) can be rewritten as

$$W(r)\,dr = \exp[-(\Delta\omega_0/\Delta\omega)^{3/p}]\,d(\Delta\omega_0/\Delta\omega)^{3/p} \tag{9-85}$$

For the case of *linear Stark effect*, the perturbing field is $F = (e/r^2)$. If we define the *normal field strength* to be

$$F_0 = (e/r_0^2) = e\left(\frac{4}{3}\pi N\right)^{\frac{2}{3}} = 2.5985eN^{\frac{2}{3}} \tag{9-86}$$

and measure F in units of F_0 [i.e., $\beta \equiv (F/F_0)$], then nearest-neighbor theory yields

$$W(\beta)\,d\beta = \frac{3}{2}\beta^{-\frac{5}{2}}\exp(\beta^{-\frac{3}{2}})\,d\beta \tag{9-87}$$

Clearly, as $\beta \to \infty$, $W(\beta) \to \frac{3}{2}\beta^{-\frac{5}{2}}$; hence statistical theory predicts that, in the wings of a line broadened by linear Stark effect, the profile falls off as $\Delta\omega^{-\frac{5}{2}}$, 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

$$W(\beta) = (2\beta/\pi)\int_0^\infty \exp(-y^{3/p})y\sin\beta y\,dy \tag{9-88}$$

Here $\beta = F/F_0$, the normal field strength F_0 now being defined as

$$F_0 = \gamma C_p N^{p/3} \tag{9-89}$$

where

$$\gamma = \{(2\pi^2 p)/[3(p+3)\Gamma(3/p)\sin(3\pi/2p)]\}^{p/3} \tag{9-90}$$

In particular, for linear Stark effect, $p = 2$, $C_p = e$, and $\gamma = 2.6031$ so that $F_0 = 2.6031eN^{\frac{2}{3}}$, 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

$$W(\beta) = \left(\frac{4}{3\pi}\right) \sum_{l=0}^{\infty} (-1)^l \Gamma\left(\frac{4l+6}{3}\right) \frac{\beta^{(2l+2)}}{(2l+1)!} \tag{9-91}$$

while for $\beta \quad 1$ one finds an asymptotic expansion

$$W(\beta) = 1.496\beta^{-\frac{5}{2}}(1 + 5.107\beta^{-\frac{3}{2}} + 14.43\beta^{-3} + \cdots) \tag{9-92}$$

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

$$n_e W_e\,dV = n_e \exp(\psi)\,dV \approx n_e(1 + \psi)\,dV \tag{9-93}$$

and

$$n_i W_i\,dV = n_i \exp(-Z_i\psi)\,dV \approx n_i(1 - Z_i\psi)\,dV \tag{9-94}$$

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,

$$n_e = \sum_i Z_i n_i \tag{9-95}$$

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*

$$\nabla^2\phi = -4\pi e\rho \tag{9-96}$$

to determine ϕ, where ρ is given by

$$e\rho = -en_e W_e + e\sum_i Z_i n_i W_i \tag{9-97}$$

In view of equations (9-93) through (9-95), this expression reduces to

$$e\rho = -e\psi\left(n_e + \sum_i Z_i^2 n_i\right) \tag{9-98}$$

Substituting equation (9-98) into (9-96), we may rewrite Poisson's equation as $\nabla^2\phi = (\phi/D^2)$ where

$$D \equiv (kT/4\pi e^2)^{\frac{1}{2}}\left[n_e + \sum_i Z_i^2 n_i\right]^{-\frac{1}{2}} \tag{9-99}$$

is the *Debye length*. Solving for ϕ we find $\phi = r^{-1}[Ae^{-r/D} + Be^{r/D}]$. Demanding that $\phi \to 0$ as $r \to \infty$, we set $B \equiv 0$. Further, to recover the potential of the ion itself as $r \to 0$, we set $A = Z_i e$. Thus the potential produced by an ion imbedded in a plasma is

$$\phi(r) = [Z_i e \exp(-r/D)]/r \tag{9-100}$$

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

$$D = 4.8(T/n_e)^{\frac{1}{2}} \text{ cm} \tag{9-101}$$

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 $r \leqslant D$, but is identically zero for $r > D$. One then finds

$$W(\beta, \delta) = (2\beta\,\delta^4/\pi)\int_0^\infty e^{-\delta g(y)} y \sin(\delta^3\beta y)\,dy \tag{9-102}$$

where

$$g(y) \equiv \frac{3}{2} y^{\frac{2}{3}} \int_y^\infty (1 - z^{-1}\sin z) z^{-\frac{5}{2}}\,dz \tag{9-103}$$

and $\delta \equiv \frac{4}{3}\pi D^3 N$, the number of particles contained in the Debye sphere. As $\delta \to \infty$, one expects to recover the Holtsmark distribution; i.e., $W(\beta, \infty) = W_H(\beta)$. From equation (9-102) one can show that for $\delta \to \infty$, $W(\beta, \delta) = W_H(\beta) + \delta^{-\frac{1}{2}}F(\beta)$ where $F(\beta)$ is a bounded definite integral. Recovery of the Holtsmark distribution for large δ can also be seen from the asymptotic expansion

$$W(\beta, \delta) = 1.496\beta^{-\frac{5}{2}}(1 + 5.107\beta^{-\frac{3}{2}} - 6.12\delta^{-\frac{1}{2}}\beta^{-2} + \cdots) \quad (9\text{-}104)$$

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.

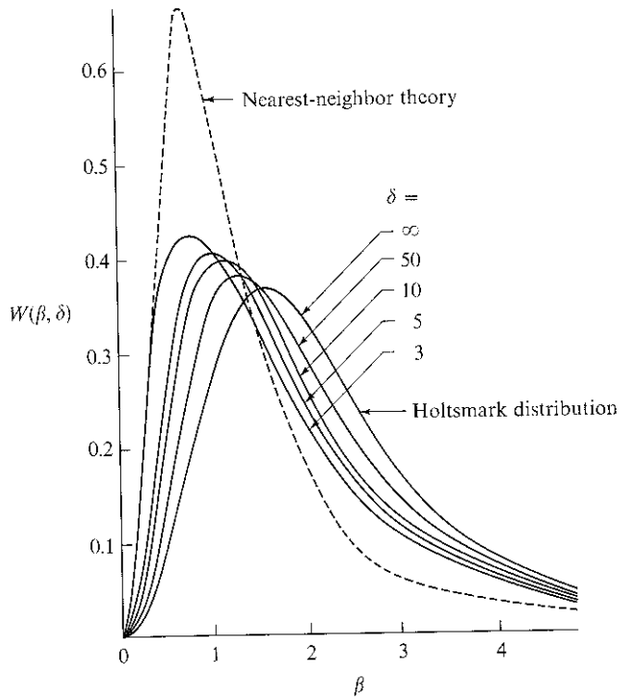

FIGURE 9-1
Probability distribution of field strength at a test point, including shielding effects; δ is the number of charged particles within the Debye sphere. From (**205**), by permission.

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

$$\Delta E = (Ze^2/r)[\exp(-r/D) - 1] \approx -Ze^2/D \quad (9\text{-}105)$$

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

$$\Delta\chi = Ze^2/D = (27.2Za_0/D)\,\text{eV} = 3 \times 10^{-8} Z n_e^{\frac{1}{2}} T^{-\frac{1}{2}}\,\text{eV} \quad (9\text{-}106)$$

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

$$\Delta\lambda_k = (3h\lambda^2 n_k/8\pi^2 cmeZ)F = C_k F \quad (9\text{-}107)$$

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

$$I(\Delta\lambda)\, d(\Delta\lambda) = \sum_k I_k W(F/F_0)(dF/F_0)$$
$$= \sum_k I_k W(\Delta\lambda/C_k F_0)\, d(\Delta\lambda)/C_k F_0 \qquad (9\text{-}108)$$

It is customary to define the parameter α as

$$\alpha \equiv \Delta\lambda/F_0 \qquad (9\text{-}109)$$

where F_0 is the normal field strength $F_0 = 2.60eN^{\frac{2}{3}}$. Then the line profile is given by a function $S(\alpha)$,

$$S(\alpha)\, d\alpha = \sum_k I_k W(\alpha/C_k)(d\alpha/C_k) \qquad (9\text{-}110)$$

which is normalized on the range $(-\infty, \infty)$ for α. The absorption cross-section per atom can be written as

$$\alpha_\nu(\Delta\lambda) = (\pi e^2/mc) f S(\Delta\lambda/F_0)(\lambda^2/cF_0) \qquad (9\text{-}111)$$

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 $\beta \gg 1$, $W(\beta) \propto \beta^{-\frac{5}{2}}$, 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 $\bar{\Delta}\omega = \bar{C}_2/r^2$ and $F = e/r^2$, it follows from equation (9-107) that

$$\bar{C}_2 = (3h\bar{n}_k/4\pi m) = 1.738\bar{n}_k \qquad (9\text{-}112)$$

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 $p = 2$)

$$\Delta\lambda_W = (v^2\lambda^2/2\pi^3 c\bar{C}_2) \qquad (9\text{-}113)$$

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 $\Delta\lambda_W$(Å) between Statistical and Impact Broadening for Hydrogen Lines

		T	
Line	Perturber	2.5×10^4 °K	10^4 °K
Hα	Electrons	580.0	230.0
	Protons	0.63	0.25
Hβ	Electrons	120.0	48.0
	Protons	0.13	0.05
Hγ	Electrons	48.0	19.0
	Protons	0.05	0.02
Hδ	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)]

$$P = (4\omega^4/3c^3)|\langle i|\mathbf{d}|j\rangle|^2 \qquad (9\text{-}114)$$

where **d** is the atomic dipole moment. The total emission, summed over all possible substates contributing to a line is

$$P(\omega) = (4\omega^4/3c^3)\sum_{i,j}\rho_j\delta(\omega - \omega_{ij})|\langle i|\mathbf{d}|j\rangle|^2 \tag{9-115}$$

Here ρ_j is the probability that an atom is in the upper state j; in thermodynamic equilibrium

$$\rho_j = \langle j|\rho|j\rangle = \exp(-E_j/kT)/U(T) \tag{9-116}$$

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

$$I(\omega) = \sum_{i,j}\rho_j\delta(\omega - \omega_{ij})|\langle i|\mathbf{d}|j\rangle|^2 \tag{9-117}$$

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

$$e^{-H/kT} = e^{-(H_A+H_P+V)/kT} \tag{9-118}$$

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

$$\phi(t) = \int_{-\infty}^{\infty} I(\omega)e^{i\omega t}\,d\omega = \sum_{i,j}\rho_j e^{i\omega_{ij}t}|\langle i|\mathbf{d}|j\rangle|^2 \tag{9-119}$$

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

$$|\alpha, t\rangle = T(t, 0)|\alpha, 0\rangle \tag{9-120}$$

Now $|\alpha, t\rangle$ satisfies the Schrödinger equation

$$H|\alpha, t\rangle = i\hbar(d|\alpha, t\rangle/dt) \tag{9-121}$$

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

$$HT(t, 0) = i\hbar[dT(t, 0)/dt] \tag{9-122}$$

which has the solution

$$T(t, 0) = e^{-iHt/\hbar} \tag{9-123}$$

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

$$\begin{aligned}\phi(t) &= \sum_{i,j}\rho_j e^{i(E_j-E_i)t/\hbar}|\langle i|\mathbf{d}|j\rangle|^2 \\ &= \sum_{i,j}\rho_j\langle j|\mathbf{d}|i\rangle e^{-iE_it/\hbar}\langle i|\mathbf{d}|j\rangle e^{iE_jt/\hbar} \\ &= \sum_{i,j}\rho_j\langle j|\mathbf{d}T|i\rangle\langle i|\mathbf{d}T^\dagger|j\rangle\end{aligned} \tag{9-124}$$

But the *expansion rule*, relative to a *complete set* of states $|\gamma\rangle$, is

$$\langle\alpha|\beta\rangle = \sum_\gamma\langle\alpha|\gamma\rangle\langle\gamma|\beta\rangle \tag{9-125}$$

so we see that equation (9-124) can be rewritten as

$$\phi(t) = \sum_j\rho_j\langle j|\mathbf{d}T\,\mathbf{d}T^\dagger|j\rangle = \mathrm{tr}(\rho\,\mathbf{d}T\,\mathbf{d}T^\dagger) \tag{9-126}$$

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

$$i\hbar(d\psi/dt) = (H_A + H_P + V)\psi \tag{9-127}$$

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 $mv\rho \gg \hbar$. But $mv\rho$ is just the angular momentum of the perturber ($= l\hbar$); 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

$$i\hbar[d\pi(t)/dt] = H_P\pi(t) \tag{9-128}$$

and the time-development operator for the perturber is

$$T_P(t, 0) = e^{-iH_P t/\hbar} \tag{9-129}$$

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

$$i\hbar[d\alpha(t)/dt] = \left(H_A + \int \pi^* V\pi \, d\tau_P\right)\alpha(t) \tag{9-130}$$

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

$$\int (\pi^* V\pi)\, d\tau_P \to V_{cl}(t) \tag{9-131}$$

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

$$i\hbar[dT_A(t, 0)/dt] = [H_A + V_{cl}(t)]T_A(t, 0) \tag{9-132}$$

and the time-development operator for the complete system is

$$T(t, 0) = T_A(t, 0)T_P(t, 0) = T_A(t, 0)e^{-iH_P t/\hbar} \tag{9-133}$$

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

$$\phi(t) = \mathrm{tr}\{\rho_A \mathbf{d} T_A \mathbf{d} T_A^\dagger\} \tag{9-134}$$

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 $\langle\alpha|\mathbf{d}|\alpha'\rangle = 0$, $\langle\beta|\mathbf{d}|\beta'\rangle = 0$, and $\langle\alpha|T_{a,b}|\beta\rangle = 0$. Then writing out the trace in equation (9-134) we find

$$\phi(t) = \rho_a \sum_{\alpha,\alpha',\beta,\beta'} \langle\alpha|\mathbf{d}|\beta\rangle\langle\beta'|\mathbf{d}|\alpha'\rangle\langle\alpha|\langle\beta|\{T_b T_a^*\}|\alpha'\rangle|\beta'\rangle \tag{9-135}$$

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

$$U(t, 0) = e^{iH_A t/\hbar} T(t, 0) \tag{9-136}$$

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

$$\phi(t) = \rho_a \sum_{\alpha, \alpha', \beta, \beta'} \langle\alpha|\mathbf{d}|\beta\rangle\langle\beta'|\mathbf{d}|\alpha'\rangle\langle\alpha|\langle\beta|e^{-iH_b t/\hbar}\{U_b U_a^*\}e^{iH_a t/\hbar}|\alpha'\rangle|\beta'\rangle \tag{9-137}$$

Substituting equation (9-136) into (9-132) we obtain a Schrödinger equation for $U(t, 0)$,

$$i\hbar[dU(t, 0)/dt] = e^{iH_A t/\hbar} V_{cl}(t) e^{-iH_A t/\hbar} U(t, 0) \equiv V'_{cl}(t) U(t, 0) \tag{9-138}$$

We solve this equation by *iteration*, starting with $U(t, 0) = 1$ as a first approximation, to obtain a solution of the form

$$U(t, 0) = 1 + (i\hbar)^{-1} \int_0^t V'_{cl}(t_1)\, dt_1 + (i\hbar)^{-2} \int_0^t dt_2 V'_{cl}(t_2) \int_0^{t_2} dt_1 V'_{cl}(t_1) + \cdots \tag{9-139}$$

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

$$\{U_b(t + \Delta t) U_a^*(t + \Delta t) - 1\} = e^{i(H_b - H_a)t/\hbar}(\Phi_{ab}\, \Delta t) e^{-i(H_b - H_a)t/\hbar} \tag{9-140}$$

We thus obtain a differential equation for $\{U_b U_a^*\}$, namely

$$\Delta\{U_b U_a^*\} = e^{i(H_b - H_a)t/\hbar} \Phi_{ab} e^{-i(H_b - H_a)t/\hbar} \{U_b U_a^*\}\, \Delta t \tag{9-141}$$

whose solution is

$$\{U_b(t, 0) U_a^*(t, 0)\} = e^{i(H_b - H_a)t/\hbar} \exp[-i(H_b - H_a)t/\hbar + \Phi_{ab} t] \tag{9-142}$$

Now, substituting into equation (9-137), we obtain

$$\phi(t) = \rho_a \sum_{\alpha, \alpha', \beta, \beta'} \langle\alpha|\mathbf{d}|\beta\rangle\langle\beta'|\mathbf{d}|\alpha'\rangle\langle\alpha|\langle\beta|e^{[i(H_b - H_a)t/\hbar + \Phi_{ab} t]}|\alpha'\rangle|\beta'\rangle \tag{9-143}$$

Then performing the inverse Fourier transformation, and reintroducing the quasi-static ion field F, we obtain the intensity profile

$$I(\omega, F) = \frac{\rho_a}{\pi} \mathscr{R}e \left(\sum_{\alpha, \alpha', \beta, \beta'} \langle\alpha|\mathbf{d}|\beta\rangle\langle\beta'|\mathbf{d}|\alpha'\rangle \times \langle\alpha|\langle\beta|\left[i\omega - \Phi_{ab}(F) - \frac{i}{\hbar}[H_a(F) - H_b(F)] \right]^{-1} |\alpha'\rangle|\beta'\rangle \right) \tag{9-144}$$

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

$$I(\omega) = \frac{\rho_a}{\pi} \int_0^\infty W(F)\, \mathscr{R}e \left(\sum_{\alpha, \alpha', \beta, \beta'} \langle\alpha|\mathbf{d}|\beta\rangle\langle\beta'|\mathbf{d}|\alpha'\rangle \times \langle\alpha|\langle\beta|\left[i\omega - \Phi_{ab}(F) - \frac{i}{\hbar}[H_a(F) - H_b(F)] \right]^{-1} |\alpha'\rangle|\beta'\rangle \right) dF \tag{9-145}$$

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

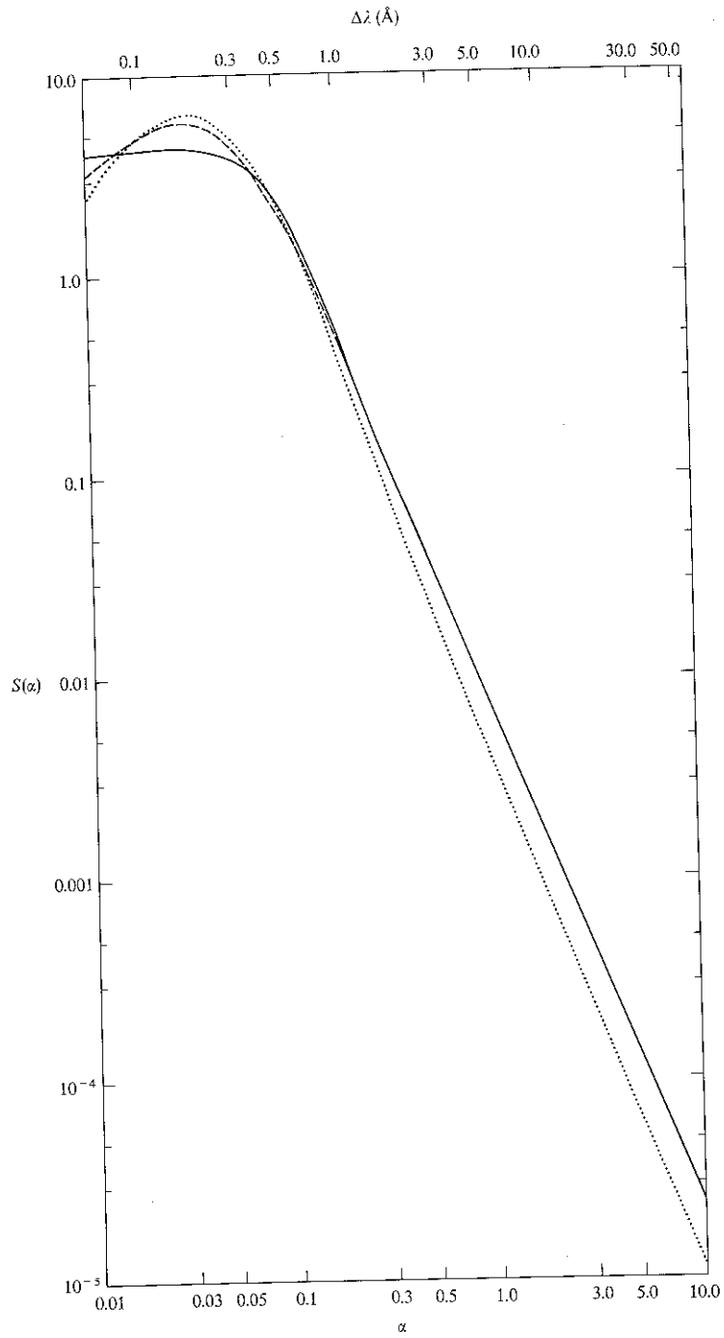

FIGURE 9-2
Stark profiles for $H\alpha$ at $n_e = 3.16 \times 10^{14}$ cm^{-3}, $T = 10^4$ °K. *Dotted curve*: Holtsmark profile for ion broadening only (**634**). *Dashed curve*: Profile for quasi-static ion broadening and impact broadening by electrons (**650**). *Solid curve*: Electron and ion-broadened profile convolved with thermal Doppler distribution of atoms. Notice that the line-core is dominated by Doppler effects, and that the absorption cross-section in the line wing is substantially larger when electron broadening is included.

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 $a = (\Gamma_{\text{rad}} + \Gamma_{\text{res}})/4\pi\,\Delta\nu_D$, and $v = (\Delta\nu/\Delta\nu_D)$. This Voigt profile is then convolved (numerically) with the Stark profile $S(\alpha)$, yielding the cross-section

$$\alpha_\nu(\Delta\nu) = (\pi^{\frac{1}{2}}e^2/mc)f \int_{-\infty}^{\infty} S^*(\Delta\nu + v\,\Delta\nu_D)H(a, v)\,dv \qquad (9\text{-}146)$$

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

$$S(\alpha) = Z^5 S_{\text{QS}}(Z^5\alpha) \qquad (9\text{-}147)$$

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 λ3203, λ4686 (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 λ304 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

$$I(\Delta\omega) = \left(\frac{w}{\pi}\right) \int_0^\infty \frac{W(F)\, dF}{(\Delta\omega - d - C_4F^2/e^2)^2 + w^2} \qquad (9\text{-}148)$$

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 $n \geqslant 4$. As was first recognized by Struve (**614**; **615**), the (2^3P–4^3D) He I λ4471 line shows a "*forbidden*" *component* (2^3P–4^3F) at λ4470. 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 λ4471 (**68**) and λ4921 (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**).

10

Classical Treatments of Line Transfer

In this chapter we discuss some of the early approaches to the line-formation problem; these provide background for the more modern treatments to be presented in following chapters. A fuller discussion of these older methods and of their application to stellar spectra can be found in (**684**, Chap. 7; **638**, Chaps. 15–17; **15**, Chaps. 12–16; **11**, Chap. 8; **256**, Chaps. 14–16). It is essential to be familiar with such methods because of the large body of literature based upon them. More important, one must understand the physical basis of the classical treatments in order to evaluate the reliability of spectroscopic diagnostics derived from them, and to realize the conceptual differences inherent in recent work.

10-1 Characterization of the Problem

In the usual classical approach one notes, at the outset, the existence of two different line-formation processes: scattering and absorption. We have discussed these two categories in Chapter 2, and described there the physical differences between them. It is usually supposed that a fraction $(1 - \varepsilon)$ of the

photons absorbed are simply *scattered* so that the excited electron returns directly to its original lower level, and normally it is assumed that the scattering is *isotropic* and *coherent* (actually *complete redistribution* is a much better approximation). Then the scattering contribution to the emission coefficient is

$$\eta_\nu^s = (1 - \varepsilon)\chi_l\phi_\nu J_\nu \tag{10-1}$$

where

$$\chi_l = (\pi e^2/mc) f_{ij}[n_i - (g_i/g_j)n_j] \tag{10-2}$$

is the line absorption coefficient in the transition between levels i and j. The remaining fraction ε of the photons is assumed to be *destroyed* and converted to thermal energy by various processes (cf. §2-1). One then argues that this loss into the thermal pool must be balanced exactly by *thermal emission*, which contributes to the total emission coefficient an amount

$$\eta_\nu^t = \varepsilon\chi_l\phi_\nu B_\nu(T) \tag{10-3}$$

In the limit of strict LTE, $\varepsilon \equiv 1$, and all the emission is thermal.

In addition, there are contributions to the opacity and emissivity from continuum thermal processes and electron scattering. Thus the transfer equation is

$$\mu(\partial I_\nu/\partial z) = -(\kappa_c + \sigma + \chi_l\phi_\nu)I_\nu + \kappa_c B_\nu + \sigma J_\nu + \varepsilon\chi_l\phi_\nu B_\nu + (1 - \varepsilon)\chi_l\phi_\nu J_\nu \tag{10-4}$$

We have ignored the ν-dependence of κ_c and σ because most lines are so narrow that these coefficients vary only negligibly over the line in comparison with the swift variation of ϕ_ν. If we write $d\tau_\nu \equiv -(\kappa_c + \sigma + \chi_l\phi_\nu)\,dz$, and let $\rho \equiv \sigma/(\kappa_c + \sigma)$, and

$$\beta_\nu \equiv \chi_l\phi_\nu/(\kappa_c + \sigma) \tag{10-5}$$

we have

$$\mu(\partial I_\nu/\partial\tau_\nu) = I_\nu - \{[(1 - \rho) + \varepsilon\beta_\nu]B_\nu + [\rho + (1 - \varepsilon)\beta_\nu]J_\nu\}/(1 + \beta_\nu) \tag{10-6}$$

Or, defining

$$\lambda_\nu \equiv [(1 - \rho) + \varepsilon\beta_\nu]/(1 + \beta_\nu) \tag{10-7}$$

the transfer equation becomes

$$\mu(\partial I_\nu/\partial\tau_\nu) = I_\nu - \lambda_\nu B_\nu - (1 - \lambda_\nu)J_\nu \tag{10-8}$$

E. A. Milne and A. S. Eddington developed equation (10-8) as an approximation to the line-transfer problem, and customarily it bears their name.

Excellent discussions of the physical implications of this equation have been given by Milne (**416**, p. 169ff) and Strömgren (**613**).

From a physical point of view, equation (10-8) provides a rather severe idealization of line formation and may be criticized on several counts. (1) Line-scattering is not, in fact, coherent. This inadequacy is overcome if we write, instead of equation (10-8),

$$\mu\left(\frac{\partial I_\nu}{\partial \tau_\nu}\right) = I_\nu - \lambda_\nu B_\nu - \frac{\rho J_\nu}{1+\beta_\nu} - \frac{(1-\varepsilon)\beta_\nu}{(1+\beta_\nu)\phi_\nu}\int R(\nu', \nu) J_{\nu'}\, d\nu' \quad (10\text{-}9)$$

where $R(\nu', \nu)$ is an appropriate redistribution function (cf. §2-1 and Chap. 13). (2) To solve equations (10-8) or (10-9), both the parameter ε and the occupation numbers n_i and n_j must be known. In classical treatments it is often assumed that LTE holds, so that $\varepsilon = 1$ and $n_i = n_i^*$, $n_j = n_j^*$. It must be stressed, however, that this merely an *assumption* and, as will be seen in Chapters 11 and 12, the assumption is often unjustified and may yield results seriously in error. In many treatments, an ad hoc value is chosen for ε, but the occupation numbers are still assumed to have their LTE values. Such an approach is internally *inconsistent*, for when scattering occurs in a line, the level-populations depend upon the radiation field via the equations of statistical equilibrium. (3) An equation of the form of (10-8) was derived by Milne for a strict two-level atom, and his analysis yields a unique (and correct) value for the parameter ε (**416**, pp. 172–178). However, as mentioned above, coherent scattering is *not* an accurate approximation. More important, analysis of the equations of statistical equilibrium for general (i.e., multilevel) atomic models shows (cf. Chapters 11 and 12) that other *kinds* of terms appear in the source function; these may depend upon the radiation fields in *other* transitions (continua and lines), and thus in principle couple *all* the lines in the spectrum together. In short, both equations (10-8) and (10-9) are seriously incomplete from a physical point of view, and this should be borne in mind during the discussion that follows in this chapter.

10-2 The Milne–Eddington Model

DEFINITION

Let us now consider the Milne–Eddington equation [equation (10-8)] under the simplifying assumptions that λ_ν, ε, and ρ are all constant with depth, and that the Planck function B_ν is a linear function on the continuum optical depth scale τ; i.e.,

$$B_\nu = a + b\tau = a + [b\tau_\nu/(1+\beta_\nu)] \equiv a + p_\nu\tau_\nu \quad (10\text{-}10)$$

Under these conditions an *exact* solution may be obtained (**158**) but this differs only slightly from the approximate solution derived below.

Taking the zero-order moment of equation (10-8) we find

$$(dH_\nu/d\tau_\nu) = J_\nu - (1-\lambda_\nu)J_\nu - \lambda_\nu B_\nu = \lambda_\nu(J_\nu - B_\nu) \quad (10\text{-}11)$$

Aside from the definition of λ_ν, equation (10-11) is the same as equation (6-4), and from the analysis in §6-1 we know that the solution, in the Eddington approximation, is

$$J_\nu = a + p_\nu\tau_\nu + (p_\nu - \sqrt{3}a)\exp[-(3\lambda_\nu)^{\frac{1}{2}}\tau_\nu]/[\sqrt{3} + (3\lambda_\nu)^{\frac{1}{2}}] \quad (10\text{-}12)$$

and the emergent flux is

$$H_\nu(0) = (3)^{-\frac{1}{2}}J_\nu(0) = \frac{1}{3}[p_\nu + (3\lambda_\nu)^{\frac{1}{2}}a]/(1+\lambda_\nu^{\frac{1}{2}}) \quad (10\text{-}13)$$

Equation (10-12) shows that *thermalization* ($J_\nu \to B_\nu$) occurs only at depths of the order of $\lambda_\nu^{-\frac{1}{2}}$. Recalling the definition of λ_ν, we see that this depth is $(1-\rho)^{-\frac{1}{2}}$ in the continuum ($\beta_\nu = 0$), and $\varepsilon^{-\frac{1}{2}}$ in a strong line ($\beta_\nu \to \infty$). In both cases thermalization occurs at a depth $p^{-\frac{1}{2}}$, where p is the probability that a photon is destroyed and converted to thermal energy each time it interacts with the material; these results are compatible with the random-walk arguments given in Chapter 6. Note that these results apply only for *coherent* scattering (cf. Chapter 11).

Equation (10-13) may be used to compute the profile of a line in a stellar atmosphere. In the continuum, $\beta_\nu = 0$; then $\lambda_\nu = (1-\rho)$, and the *continuum flux* is

$$H_c(0) = \frac{1}{3}[b + a\sqrt{3(1-\rho)}]/[1 + (1-\rho)^{\frac{1}{2}}] \quad (10\text{-}14)$$

Thus the *residual flux* in the line is, by equation (8-2),

$$R_\nu = \left[\frac{p_\nu + (3\lambda_\nu)^{\frac{1}{2}}a}{1+\lambda_\nu^{\frac{1}{2}}}\right]\left[\frac{1+(1-\rho)^{\frac{1}{2}}}{b + a\sqrt{3(1-\rho)}}\right] \quad (10\text{-}15)$$

There are four important results that can be derived from this classical theory, which provided the conceptual orientation of much of the early work on line formation. Let us now examine these briefly.

SCATTERING LINES

Consider the case where $\rho = 0$, so that there is no scattering in the continuum. Further, suppose that $\varepsilon = 0$, so that there is pure scattering in the

lines. Then $\lambda_\nu = (1 + \beta_\nu)^{-1}$ and the residual flux in the line is

$$R_\nu = 2\left[\frac{b}{1+\beta_\nu} + a\left(\frac{3}{1+\beta_\nu}\right)^{\frac{1}{2}}\right]\left[\left(1 + \sqrt{\frac{1}{1+\beta_\nu}}\right)(\sqrt{3}a + b)\right]^{-1} \tag{10-16}$$

If we consider the case of a very strong line and take the limit $\beta_\nu \to \infty$, we obtain $R_\nu = H_\nu(0)/H_c(0) = 0$, which shows that the core of a very strong line formed by scattering can be completely dark; this result is in contrast to the case of a line formed by absorption, as is shown below.

Exercise 10-1: Writing $\beta_\nu = \beta_0 H(a, v)$, where H denotes the usual Voigt function, plot residual flux profiles from equation (10-16) for scattering lines with $\beta_0 = 1, 10, 100, 1000, 10^4$, and $(b/a) = 1, 2$, and 3, assuming $a = 10^{-3}$.

Exercise 10-2: In the *Schuster–Schwarzschild model*, the lines are assumed to be confined to a finite layer (the "*reversing layer*") of thickness τ_ν, illuminated from below by an incident intensity I_0. In the reversing layer the continuum opacity is zero, and the lines are pure scatterers. Using the *two-stream approximation* ($I \equiv I^+$ for $0 \leqslant \mu \leqslant 1$; $I \equiv I^-$ for $-1 \leqslant \mu \leqslant 0$; and $\mu = \pm\frac{1}{2}$ in the transfer equation) show that (a) $H_\nu \equiv \frac{1}{2}(I_\nu^+ - I_\nu^-) = \text{constant} = \frac{1}{2}I_0/(1 + \tau_\nu)$, and (b) $J_\nu(t_\nu) = H_\nu(1 + 2t_\nu)$, $0 \leqslant t_\nu \leqslant \tau_\nu$.

ABSORPTION LINES

Again assume $\rho = 0$, but now set $\varepsilon = 1$ (LTE in the line); then $\lambda_\nu \equiv 1$, and

$$R_\nu = [\sqrt{3}a + b(1 + \beta_\nu)^{-1}]/(\sqrt{3}a + b) \tag{10-17}$$

In this case, as $\beta_\nu \to \infty$, the flux in the line does *not* go to zero, but approaches a finite value

$$R_0 = R_\nu(\varepsilon = 1, \quad \beta_\nu \to \infty) = [1 + (b/\sqrt{3}a)]^{-1} \tag{10-18}$$

This is to be expected, for as $\beta_\nu \to \infty$, only the surface layers of the star are seen, and the emergent flux is then determined by the surface value of the Planck function, which will be nonzero. In contrast, in the scattering case, photons are constantly diverted out of the pencil of radiation, and in the limit $\beta_\nu \to \infty$, none survive to emerge at the surface.

Exercise 10-3: Repeat exercise 10-1 for an absorption line, using equation (10-17).

It is convenient to re-express equations (10-17) and (10-18) in terms of the Planck function and its gradient. Assume that on a *mean* optical depth scale,

$$B_\nu(\bar{\tau}) = B_\nu(T_0) + (\partial B_\nu/\partial\bar{\tau})_0\bar{\tau} \equiv B_0 + B_1\bar{\tau} \tag{10-19}$$

Using the Eddington-approximation result for the grey temperature distribution, namely $T^4 = T_0^4(1 + \frac{3}{2}\bar{\tau})$, it is easy to show that

$$B_1 = \frac{3}{8}X_0B_0 \tag{10-20}$$

where

$$X_0 \equiv u_0/(1 - e^{-u_0}) \tag{10-21}$$

and $u_0 \equiv (h\nu/kT_0)$. Thus the parameters in equation (10-10) are: $a \equiv B_0$,

$$b = \frac{3}{8}X_0B_0(\bar{\kappa}/\kappa) \tag{10-22}$$

and

$$p_\nu = \frac{3}{8}X_0B_0(\bar{\kappa}/\kappa)/(1 + \beta_\nu) \tag{10-23}$$

where $\bar{\kappa}$ and κ are the mean opacity, and the monochromatic continuum opacity at the line wavelength, respectively. Then equation (10-18) becomes

$$R_0 = \{1 + [\sqrt{3}X_0(\bar{\kappa}/\kappa)/8]\}^{-1} \tag{10-24}$$

Exercise 10-4: Derive equations (10-20) through (10-23).

For the sun, $T_0 \sim 4800°$K from a grey model, and if we choose $\lambda = 5000$ Å, then $u_0 \sim 6$, $X_0 \sim 6$, and $\kappa \approx \bar{\kappa}$, so that equation (10-24) predicts

$$R_0 = \left(1 + \frac{3}{4}\sqrt{3}\right)^{-1} = 0.44 \tag{10-25}$$

This value is in fair agreement with the depths of many of the stronger lines observed in that region of the solar spectrum. Some lines, however, are much deeper, particularly resonance lines such as the sodium D-lines; this fact led to the conceptual identification of resonance lines as "scattering" lines and subordinate lines (e.g., $H\alpha$) as "absorption" lines. It was believed that the central intensities of the latter group reflected information about the surface temperature of the atmosphere. Such a division of lines into two groups is intuitively not unreasonable, for we expect that in a resonance line the most probable route of exit from the upper state is, in fact, a direct decay to the lower state. In contrast, for subordinate lines a large number of possibilities will, in general, exist, and the photons may effectively be removed from the line and destroyed. It must be emphasized, however, that this characterization is only schematic and often does violence to the actual physics of line-formation. For example, we shall find in Chapter 11 that the boundary value

of the source function of the $H\alpha$ line (a subordinate line) has essentially nothing to do with the temperature of the outer atmospheric layers.

CENTER-TO-LIMB VARIATION

The specific intensity emergent at frequency ν, on the disk at an angle $\theta = \cos^{-1}\mu$ from disk center, is given by

$$I_\nu(0,\mu) = \int_0^\infty S_\nu(\tau_\nu) e^{-(\tau_\nu/\mu)} \mu^{-1}\, d\tau_\nu$$

$$= \int_0^\infty [B_\nu + (1-\lambda_\nu)(J_\nu - B_\nu)] \exp(-\tau_\nu/\mu)\mu^{-1}\, d\tau_\nu \quad (10\text{-}26)$$

where the form of S_ν has been taken from equation (10-8). Substituting from equation (10-12) we find

$$I_\nu(0,\mu) = (a + p_\nu\mu) + \left[\frac{(p_\nu - \sqrt{3}a)(1-\lambda_\nu)}{\sqrt{3}(1+\sqrt{\lambda_\nu})(1+\sqrt{3\lambda_\nu}\mu)}\right] \quad (10\text{-}27)$$

In the continuum, $\beta_\nu = 0$, and again taking $\rho = 0$,

$$I_c(0,\mu) = a + b\mu \quad (10\text{-}28)$$

The residual intensity $r_\nu(\mu) \equiv I_\nu(0,\mu)/I_c(0,\mu)$ and the absorption depth $a_\nu(\mu) = 1 - r_\nu(\mu)$ follow immediately from equations (10-27) and (10-28).

Consider first a pure absorption line, for which $\varepsilon = 1$ and $\lambda_\nu \equiv 1$. Then

$$r_\nu(\mu) = [1 + (b/a)\mu/(1+\beta_\nu)]/[1 + (b/a)\mu] \quad (10\text{-}29)$$

Here we see that as the limb is approached (i.e., as $\mu \to 0$), $I_\nu \to I_c$, and the line *vanishes*. This is compatible with the physical picture sketched above, for as the limb is reached, only the very surface can be seen at any frequency. For $\varepsilon = 1$, the same source function (the surface value of the Planck function) is observed throughout the profile, and the contrast between line and continuum disappears. On the other hand, if $\varepsilon = 0$ (a pure scattering line), $\lambda_\nu = (1+\beta_\nu)^{-1}$, and taking the limit as $\beta_\nu \to \infty$, equation (10-27) reduces to $I_\nu(0,\mu) \equiv 0$ for all μ's; thus the cores of pure scattering lines always remain dark, even at the limb, and there is a clear distinction between the center-to-limb behavior of absorption and scattering lines. A few representative values of the variation of line-depth with disk-position are given in Table 10-1 for several values of ε and β_ν. In this table $(b/a) = 3$; negative values imply emission, and are an artifact of the approximations made in solving the transfer equation.

TABLE 10-1
Center-to-Limb Variation in the Milne–Eddington Model

		$a_\nu(\mu)$				
ε	β_ν	$\mu = 1$	0.5	0.3	0	A_ν (Flux)
1.0	0.01	0.007	0.006	0.005	0.000	0.006
	0.1	0.068	0.055	0.043	0.000	0.058
	1.0	0.375	0.300	0.237	0.000	0.317
	10.0	0.682	0.545	0.431	0.000	0.576
	100.0	0.743	0.594	0.469	0.000	0.628
	∞	0.750	0.600	0.474	0.000	0.634
0.3	0.01	0.007	0.005	0.004	−0.002	0.006
	0.1	0.066	0.051	0.037	−0.019	0.054
	1.0	0.378	0.306	0.246	0.026	0.322
	10.0	0.723	0.633	0.565	0.334	0.653
	100.0	0.798	0.713	0.648	0.438	0.731
	∞	0.808	0.723	0.659	0.452	0.741
0.1	0.01	0.007	0.005	0.004	−0.003	0.006
	0.1	0.066	0.049	0.035	−0.024	0.053
	1.0	0.379	0.308	0.250	0.035	0.324
	10.0	0.751	0.687	0.639	0.483	0.700
	100.0	0.847	0.799	0.765	0.658	0.809
	∞	0.860	0.815	0.783	0.684	0.824
0.0	0.01	0.007	0.005	0.003	−0.004	0.006
	0.1	0.066	0.049	0.034	−0.027	0.053
	1.0	0.379	0.310	0.252	0.039	0.325
	10.0	0.778	0.732	0.698	0.589	0.742
	100.0	0.931	0.920	0.912	0.885	0.922
	∞	1.000	1.000	1.000	1.000	1.000

In the solar spectrum, some lines do weaken towards the limb while others do not, or weaken only slightly. This observed behavior again led to the classification of lines into "absorption" or "scattering" categories, though in some cases the categorization made on the basis of limb-darkening was in conflict with that based on central intensities. Furthermore, some studies [e.g., that by Houtgast, reviewed in (**596**)] have shown that *neither* category is adequate because in some lines the effects of noncoherent scattering dominate. In short, the approach described here is quite schematic, and one must recognize that it simply does not contain much of the essential physics.

SCHUSTER MECHANISM

In the above discussion, the continuum has been assumed to be purely thermal. When continuum scattering is taken into account, several interesting

effects are found; these were first discussed by Schuster in one of the fundamental papers of radiative transfer theory (**562**). To emphasize the role of continuum scattering, set $\rho = 1$ (a realistic value for, say, O-stars); then $\lambda_\nu = \varepsilon\beta_\nu/(1 + \beta_\nu)$ and equation (10-15) becomes

$$R_\nu = \left[\frac{1}{1+\beta_\nu} + \left(\frac{a}{b}\right)\left(\frac{3\varepsilon\beta_\nu}{1+\beta_\nu}\right)^{\frac{1}{2}}\right]\left[1 + \left(\frac{\varepsilon\beta_\nu}{1+\beta_\nu}\right)^{\frac{1}{2}}\right]^{-1} \tag{10-30}$$

First, consider the case with $\varepsilon = 0$; then $R_\nu = 1/(1 + \beta_\nu)$, and the line is a pure absorption feature. This occurs because both the lines and continuum are formed by scattering, and the scattering length in the line is larger. On the other hand, suppose $\varepsilon = 1$; then it is obvious from equation (10-30) that in the line core, as $\beta_\nu \to \infty$, $R_\nu \to (\sqrt{3}a/2b)$. Thus the line can appear in absorption or *emission* depending on the ratio (a/b); the more shallow the temperature gradient (i.e., the smaller the value of b), the brighter is the line compared to the continuum. The reason for this is that because $\varepsilon = 1$, the source function in the line is everywhere equal to the thermal value while, in the continuum, scattering causes J_ν (and thus S_ν) to drop below B_ν.

If (a/b) is greater than the critical value $(a/b) = 2/\sqrt{3}$, the line is in emission for all values of β_ν (i.e., at all frequencies in the profile). If (a/b) just equals the critical value, then both the extreme wing ($\beta_\nu \to 0$) and the very core of the line ($\beta_\nu \to \infty$) lie at the level of the continuum, and all other points in the profile are in emission. If (a/b) is less than the critical value the line may have weak emission wings and a central reversal into an absorption core. If $(a/b) < 1/\sqrt{3}$, the line is everywhere in absorption.

This interplay of scattering and absorption, which gives rise to either absorption or emission lines depending on the temperature gradient, is known as the *Schuster mechanism*; a thorough discussion of the various possible cases was given by Schuster himself (**562**). From time to time it has been suggested that this mechanism could be responsible for emission lines observed in certain early-type spectra, but on the basis of recent critical discussions (**238**; **280**) this seems unlikely.

10-3 The Theoretical Curve of Growth

Using equation (10-15) one could, in principle, compute the profile of a spectrum line and, by integration over frequency, determine its equivalent width. Such a procedure is laborious, however, and requires the use of a computer. It is instructive, therefore, to consider a simple model that allows the equivalent width to be computed analytically. In this way we can construct what is known as a *curve of growth*, which gives the equivalent width directly in terms of the number of absorbing atoms that produce the line.

To begin with, we assume that the line formation occurs in a layer to which we can assign a unique temperature and electron pressure. This assumption is normally valid in laboratory work, but in the stellar-atmospheres situation there are usually strong gradients in both temperature and pressure. Thus the choice of an appropriate temperature and pressure is difficult and, at best, can refer only to some ill-defined mean value in the line-forming region. Having made this choice, we can compute (assuming LTE) the populations of the atomic levels and, hence, the continuous opacity κ_c and the line opacity in a transition $(i \to j)$,

$$\chi_{ij}(\nu) = (\pi e^2/mc) f_{ij} n_i^* (1 - e^{-h\nu/kT})\phi_\nu \equiv \chi_{ij}\phi_\nu \tag{10-31}$$

The line profile is assumed to be given by the Voigt function [equation (9-34)]

$$H(a, v) = (a/\pi)\int_{-\infty}^{\infty} e^{-y^2}[(v - y)^2 + a^2]^{-1}\, dy$$

where $v \equiv \Delta\nu/\Delta\nu_D$, $a \equiv \Gamma/(4\pi\,\Delta\nu_D)$, and $\Delta\nu_D = \nu\xi_0/c$; here ξ_0 is the most probable velocity of the atoms in the material. We then write $\chi_{ij}(\nu) = \chi_0 H(a, v)$ where $\chi_0 \equiv \chi_{ij}/(\pi^{\frac{1}{2}}\,\Delta\nu_D)$. We assume that both $\Delta\nu_D$ and the parameter a are *fixed* in the region of line-formation, and employ the Milne–Eddington model in which the ratio $\beta_\nu \equiv \chi_{ij}(\nu)/\kappa_c$ is *independent of depth.* (The assumption of depth-independent β_ν is actually a fairly good approximation for some spectrum lines; e.g., Mg II λ4481 and Si II $\lambda\lambda$4128, 4131. Indeed for these lines both χ_0 and κ_c may vary over orders of magnitude in the atmosphere while their ratio remains nearly constant.) Finally, we assume the line is formed by *absorption* processes in LTE; the flux is then easily computed as

$$F_\nu = 2\int_0^\infty B_\nu[T(\tau)]E_2\left[\int_0^\tau (1 + \beta_\nu)\, dt\right](1 + \beta_\nu)\, d\tau \tag{10-32}$$

As before, we adopt $B_\nu[T(\tau)] = B_0 + B_1\tau$, where τ is the continuum optical depth. Then, having taken β_ν to be constant with depth, we have

$$\begin{aligned} F_\nu &= 2\int_0^\infty (B_0 + B_1\tau)E_2[(1 + \beta_\nu)\tau](1 + \beta_\nu)\, d\tau \\ &= B_0 + \frac{2}{3}B_1/(1 + \beta_\nu) \end{aligned} \tag{10-33}$$

From equation (10-33) the continuum flux is clearly $F_c = B_0 + \frac{2}{3}B_1$, so that the line-depth in the flux profile is

$$A_\nu = [\beta_\nu/(1 + \beta_\nu)]\left[1 + \frac{3}{2}(B_0/B_1)\right]^{-1} \tag{10-34}$$

It is convenient to work in terms of the parameter A_0, the central depth of an infinitely opaque line. Taking the limit of equation (10-34) as $\beta_\nu \to \infty$ we find

$$A_0 = \left[1 + \frac{3}{2}(B_0/B_1)\right]^{-1} \tag{10-35}$$

which is essentially the result stated in equation (10-18). Then equation (10-34) may be rewritten

$$A_\nu = A_0\beta_\nu/(1 + \beta_\nu) \tag{10-36}$$

The *equivalent width* may now be computed by integration of the line-depth over frequency to obtain

$$W_\nu = \int_0^\infty A_\nu\, dv = 2A_0\,\Delta v_D \int_0^\infty \beta(v)[1 + \beta(v)]^{-1}\, dv \tag{10-37}$$

where it has been assumed that the line is symmetrical about its center. But $\beta(v) = (\chi_0/\kappa_c)H(a, v) \equiv \beta_0 H(a, v)$, and if we define the *reduced equivalent width* $W^* \equiv W/(2A_0\,\Delta v_D)$, we have

$$W^*(a, \beta_0) = \int_0^\infty \beta_0 H(a, v)[1 + \beta_0 H(a, v)]^{-1}\, dv \tag{10-38}$$

Before we attempt to evaluate this integral, let us consider qualitatively how the line develops as more and more atoms absorb. At the start, with only a few absorbers, each will be able to remove photons from the radiation field, and the line strength should be proportional to the number of absorbing atoms. Only the Doppler core (where the opacity is highest) will contribute to the line strength; the line wings will be transparent and will not reduce the emitted flux. As yet more atoms absorb in the line, the core, at some point, becomes completely opaque, and the intensity there reaches its limiting value given by equation (10-35). Now, essentially all the photons that can be absorbed in the core already have been, and so long as the line-wings remain transparent, the addition of more absorbers does little to increase the equivalent width of the line, which is said to be *saturated*. Finally, when enough absorbers are present, the opacity in the wings becomes appreciable, and the equivalent width again increases as the contribution of the line wings grows. These different behaviors are shown in Figure 10-1. On the basis of the above discussion, we see that there are three basically distinct regimes in the curve of growth; these will be treated individually.

The Voigt profile can be represented schematically [equation (9-45)] as $H(a, v) \sim e^{-v^2} + \pi^{-\frac{1}{2}}av^{-2}$; here it is assumed that the first term applies only in the core for $v \leqslant v^*$, and the second only in the wings for $v \geqslant v^*$, where v^* is chosen as the transition point where the two terms are equal.

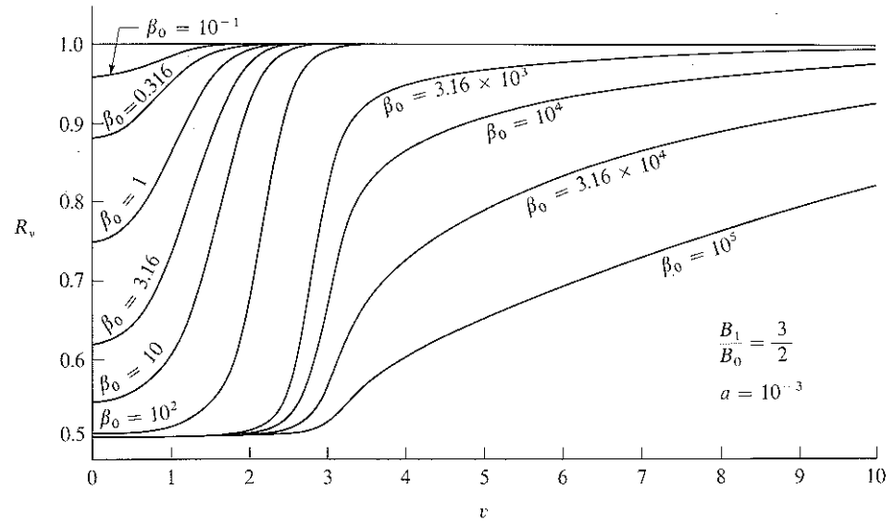

FIGURE 10-1
Development of a spectrum line with increasing number of atoms along the line of sight. The line is assumed to be formed in pure absorption. For $\beta_0 \lesssim 1$, the line strength is directly proportional to the number of absorbers. For $30 \lesssim \beta_0 \lesssim 10^3$ the line is saturated, but the wings have not yet begun to develop. For $\beta_0 \gtrsim 10^4$ the line wings are strong and contribute most of the equivalent width.

Consider now the contribution from the core only, and write $\beta(v) = \beta_0 e^{-v^2}$, assuming that $\beta_0 < 1$. Then equation (10-38) becomes

$$\begin{aligned} W^* &= \beta_0 \int_0^\infty e^{-v^2}[1 + \beta_0 e^{-v^2}]^{-1}\, dv \\ &= \beta_0 \int_0^\infty e^{-v^2}(1 - \beta_0 e^{-v^2} + \cdots)\, dv \end{aligned} \tag{10-39}$$

or

$$W^* = \frac{1}{2}\pi^{\frac{1}{2}}\beta_0[1 - (\beta_0/\sqrt{2}) + (\beta_0{}^2/\sqrt{3}) - \cdots] \tag{10-40}$$

For small values of β_0 (weak lines) the linear term dominates, and the equivalent width of the line is directly proportional to the number of absorbers present. This is known as the *linear* part of the curve of growth. Note that β_0 varies as $\Delta v_D{}^{-1}$, as does W^*, hence the equivalent width W is *independent* of Δv_D on the linear part of the curve.

In the *saturation* part of the curve of growth, β_0 is so large that the line core has reached its limiting depth, but not yet large enough that the line wings contribute to the equivalent width. Again $\beta(v) = \beta_0 e^{-v^2}$, and if we

let $u = v^2$, then $dv = \frac{1}{2}u^{-\frac{1}{2}}\,du$, and equation (10-38) becomes

$$W^* = \frac{1}{2}\int_0^\infty \beta_0 e^{-u}(1 + \beta_0 e^{-u})^{-1}u^{-\frac{1}{2}}\,du$$

$$= \frac{1}{2}\int_0^\infty (1 + e^{u-\alpha})^{-1}u^{-\frac{1}{2}}\,du \qquad (10\text{-}41)$$

where we have set $\beta_0 = e^\alpha$. This integral may be rewritten [see (**160**, 389)] as

$$2W^* = \int_0^\alpha u^{-\frac{1}{2}}\,du + \alpha\int_0^\infty \frac{[\alpha(1+t)]^{\frac{1}{2}}\,dt}{1+e^{\alpha t}} - \alpha\int_0^\alpha \frac{[\alpha(1-t)]^{\frac{1}{2}}\,dt}{1+e^{\alpha t}} \qquad (10\text{-}42)$$

Following Sommerfeld, one may replace the upper limit in the third integral by ∞ (because $\alpha \gg 1$) and expand $[\alpha(1+t)]^{\frac{1}{2}}$ as a power series in t around $t = 0$. The results may be written in closed form in terms of the Riemann zeta function, and one obtains finally the asymptotic expression

$$W^* \approx \sqrt{\ln \beta_0}\{1 - [\pi^2/24(\ln \beta_0)^2] - [7\pi^4/384(\ln \beta_0)^4] - \cdots\} \qquad (10\text{-}43)$$

The above expansion is only semiconvergent and must always be truncated after a finite number of terms; in practice, the series is useful for $\beta_0 \gtrsim 55$. From equation (10-43) we can see clearly that on the *saturation*, or *flat*, part of the curve of growth, the equivalent width grows extremely slowly with increasing numbers of absorbers, namely $W^* \propto \sqrt{\ln \beta_0}$. The weak dependence of W^* on β_0 implies that W on this part of the curve is essentially proportional to $\Delta\nu_D$. It is easy to understand why this should be so: the *depth* of the line profile is fixed at A_0, hence the integrated absorption must be proportional to the linewidth; i.e., to $\Delta\nu_D$ (cf. Figure 10-1). Similarly the dependence of W on β_0 can be understood by recognizing that to produce a significant depression in the continuum, the optical depth in the line must exceed unity before the continuum reaches unit optical depth. This occurs at frequencies $\nu < \nu_0$ where $\beta_0 e^{-\nu_0^2} \approx 1$; clearly ν_0, which measures the width of the dark core (and hence determines W^*), varies as $\sqrt{\ln \beta_0}$.

Finally, for very large numbers of absorbers, the line wings become opaque enough to provide the dominant contribution to the equivalent width. Here we adopt $H(a, \nu) \sim a/(\pi^{\frac{1}{2}}\nu^2)$, and writing $C \equiv \beta_0 a\pi^{-\frac{1}{2}}$ we find from equation (10-38)

$$W^* = \int_0^\infty (1 + \nu^2/C)^{-1}\,d\nu = \frac{1}{2}\pi\sqrt{C} = \frac{1}{2}(\pi a\beta_0)^{\frac{1}{2}} \qquad (10\text{-}44)$$

Thus $W \propto \beta_0^{\frac{1}{2}}$, giving rise to what is called the *damping* or *square-root* part of the curve of growth. Again, note that both a and β_0 each contain a factor of $\Delta\nu_D$; hence W is independent of $\Delta\nu_D$ on this part of the curve.

The entire curve of growth is shown in Figure 10-2. Notice that the larger the value of the damping parameter a, the sooner the wings dominate W,

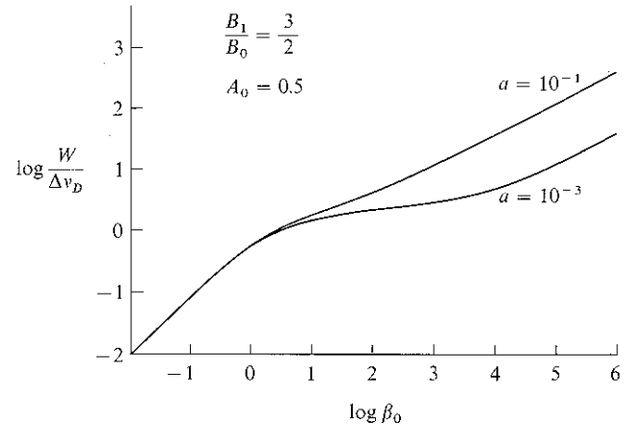

FIGURE 10-2
Curves of growth for pure absorption lines. Note that the larger the value of a, the sooner the square-root part of the curve rises away from the flat part.

and hence the sooner the damping part of the curve rises away from the flat part. Numerous curves of growth have been computed by various authors. A particularly useful set has been published by Wrubel for a wide range of the temperature-distribution parameters B_0 and B_1, and under different assumptions concerning the atmospheric model (i.e., Milne–Eddington or Schuster–Schwarzschild) and the transfer problem (i.e., absorption or scattering lines); see (**687**; **688**; **689**).

10-4 The Empirical Curve of Growth

The curve of growth has long been one of the astronomer's favorite tools for performing an analysis of a stellar atmosphere, and the literature of the various applications of this approach is enormous. The reasons for the popularity of the curve of growth are that it provides estimates of several key parameters with speed and ease, and that it makes use of equivalent widths alone, which can be well-determined observationally even for faint stars where profiles would be impossible to measure accurately.

As described in §10-3, the theoretical curve of growth gives $\log(W_\nu/\Delta\nu_D) = \log(W_\lambda/\Delta\lambda_D) = \log(W_\lambda c/\lambda\xi_0)$ as a function of $\log \beta_0$. Here ξ_0 is the total random velocity of the atoms forming the line, and

$$\beta_0 = \frac{\chi_0}{\kappa_c} = \left(\frac{\pi^{\frac{1}{2}}e^2 f}{mc\,\Delta\nu_D}\right)\frac{n_{ijk}^*(1 - e^{-h\nu/kT})}{k_c(1 - e^{-h\nu/kT})} = \left(\frac{\pi^{\frac{1}{2}}e^2}{mc}\right)(f\lambda)\left(\frac{n_{ijk}^*}{\xi_0 k_c}\right) \qquad (10\text{-}45)$$

Here k_c, the continuum opacity *uncorrected* for stimulated emissions, has been introduced for convenience (and continuum scattering has been ignored). The population n^*_{ijk}, of excitation-state i, of ionization-stage j, of chemical species k, is assumed to be given by the Boltzmann–Saha equations as described in §5-1.

In a stellar spectrum one can usually observe lines from several multiplets of a given ion, and one or more ionization stages of a given element, for each of several elements. Let us first consider only lines from a definite ion of a single element. Because of the factor $\exp(-\chi_{ijk}/kT)$ in the Boltzmann excitation formula, each multiplet has its own curve of growth. In view of equation (5-8) we may write

$$\log \beta_0 = \log(g_{ijk} f\lambda) - \theta\chi_{ijk} + \log C_{jk} \tag{10-46}$$

where $\theta \equiv 5040/T$, χ is expressed in eV, and

$$C_{jk} \equiv [N_{jk}/U_{jk}(T)](\pi^{\frac{1}{2}}e^2/mc)(\xi_0 k_c)^{-1} \tag{10-47}$$

Clearly the value of β_0 will be affected by the choice for the temperature (θ) in the atmosphere, and will be different for each multiplet (because of differences in χ_{ijk}), and for each line (because of differences in f-values).

To construct an empirical curve of growth, we plot the value of $\log(W_\lambda/\lambda)$ versus $\log(gf\lambda)$ for each line. Now if we assume that there is a *unique* relation between β_0 and W_λ, we attempt to force all points for the different lines to define a single curve as nearly as possible. To do this, we adjust θ, and choose the value that minimizes the scatter around a mean curve. This value is called the *excitation temperature*, θ_{exc}, and is considered to be the characteristic temperature of the line-forming region. It should be realized that it may not always be possible to derive a meaningful excitation temperature, for different lines will actually be formed in different regions of the atmosphere. For example, one would expect that lines from levels with high excitation potentials will be formed deeper in the atmosphere where temperatures are higher. Similarly, we would expect the average excitation temperature to be higher for higher stages of ionization. Such refinements can be taken into account by more elaborate calculations that use model atmospheres, but they are normally ignored in curve-of-growth work.

There are other complications that arise in practice: there are errors both in the values of W_λ and f, and these introduce scatter into the curve. Also, lines arising from only a limited range of excitation potentials may be observable, and θ_{exc} may not be very well determined over this short baseline.

Once the empirical curve, corrected for excitation effects, has been established, it may be compared with theoretical curves. To superimpose the two curves, a shift (both in abscissa and ordinate) of the empirical curve relative

to the theoretical is normally required. From this fitting procedure we can deduce three essential bits of information.

(1) The *ordinate* of the empirical curve is $\log(W_\lambda/\lambda)$, while that of the theoretical curve is $\log(W_\lambda/\Delta\lambda_D) = \log(W_\lambda/\lambda) - \log(\xi_0/c)$. Thus when the two curves are superimposed, *the difference in the ordinates yields* $\log(\xi_0/c)$, *and thus the velocity parameter* ξ_0. The value derived can be compared with the most probable thermal velocity at the excitation temperature T_{exc}, namely $\xi_{\text{therm}} = (2kT_{\text{exc}}/Am_{\text{H}})^{\frac{1}{2}}$ where A is the atomic weight of the element. It is usually found that ξ_0 as inferred from the curve of growth *exceeds* ξ_{therm}, sometimes by a large factor. To explain this difference, it has been customary to postulate the existence of additional nonthermal motions of the stellar material, which are usually referred to as *microturbulence*. It is assumed that these motions occur on scales that are small compared to a photon mean-free-path and hence constitute, in effect, an additional source of line-broadening. If these small-scale mass motions have a Gaussian distribution around some most probable speed ξ^2_{turb}, then

$$\xi_0 = [(2kT_{\text{exc}}/Am_{\text{H}}) + \xi^2_{\text{turb}}]^{\frac{1}{2}} \tag{10-48}$$

"Turbulent" velocities have been derived for many stars by this method of analysis; the most dramatic results are obtained for supergiants where velocities in excess of the sound speed in the gas have been obtained. It should be realized, however, that such diagnoses are not on entirely firm ground because the introduction of a velocity field drastically affects the details of line formation (cf. Chapter 14). Indeed, with the high velocities sometimes derived, one must inquire whether the excitation state of the gas is affected by interchange of energy between mass motions and the internal energy of the material. Very little work has been done on this difficult problem, and our knowledge about small-scale stellar velocity fields remains rudimentary, almost to the point of merely recognizing their existence.

(2) The difference between the *abscissae* of the empirical and theoretical curves of growth yields

$$\log C = \log \beta_0 - [\log(gf\lambda) - \theta_{\text{exc}}\chi] \tag{10-49}$$

To proceed further, we need an estimate of the electron density. This is derived from a theoretical model of the atmosphere. Once we know n_e, we may compute k_c, and as we already know ξ_0, we find N_{jk}, the number of atoms of chemical species k in ionization stage j directly from C, via equation (10-47). Then by use of the Saha ionization formula we can convert N_{jk} to N_k, the number of atoms of chemical species k in all ionization states. More precisely, we obtain the abundance $\alpha_k = (N_k/N_{\text{H}})$, as k_c is usually dominated by hydrogen or H^-, and thus is proportional to N_{H}. In short, *the horizontal*

shift between the two curves of growth yields the abundance of the element. This method of analysis has been applied to a wide variety of stars, and has shown that certain stars (e.g., Population II stars and peculiar stars) have abundances that differ markedly from solar values, which, in turn, are fairly typical of Population I stars. The uncertainties in a curve of growth abundance analysis should be clear: it has been assumed that the Saha–Boltzmann relations are valid; depth variations of the relevant parameters (called *stratification effects*) have been ignored; and a schematic solution of the transfer problem has been employed. The accuracy of the results is, therefore, limited.

(3) The horizontal and vertical shifts described above make use of the linear and flat parts of the curve respectively. For a given set of theoretical curves, a comparison between the observed and computed damping parts determines the value of $a = \Gamma/(4\pi\,\Delta\nu_D)$, and hence Γ; this in turn may be compared with the value predicted by line-broadening theory.

Let us now turn to a brief discussion of a few typical results. The solar spectrum has been analyzed extensively with curve-of-growth techniques; one of the most outstanding early treatments was by H. N. Russell (**542**), who derived element abundances (from eye-estimates of line-strengths!) that are in remarkably good agreement with current estimates. A very interesting study using curves of growth was carried out by K. O. Wright (**686**) who analyzed the Sun and three other solar-type stars. This study is by no means the most recent available [see, e.g., (**182**)], but it is a classic example of the procedure, so we shall consider it here. From an extensive set of equivalent-width measurements and laboratory f-values, an empirical curve of growth was constructed using lines of Fe I and Ti I. In all, some 75 lines of Fe I arising from states with $0 \leqslant \chi \leqslant 1.6$ eV, and 137 lines of Ti I with $0 \leqslant \chi \leqslant 2.5$ eV, were used to obtain the curve shown in Figure 10-3. Slightly different excitation temperatures were found for the two atoms, namely $T_{\text{exc}} = 4850°\text{K} \pm 150°\text{K}$ for Fe I and $T_{\text{exc}} = 4550°\text{K} \pm 150°\text{K}$ for Ti I. These values are about what we would expect for a radiative-equilibrium model atmosphere. As can be seen, the curve is well defined, though it is true that the linear part is defined mainly by Ti I lines and the damping part by Fe I lines. It would be more satisfactory if lines of each atom were found in both the linear and damping portions of the curve.

The vertical shift of the empirical curve, relative to a theoretical curve by Menzel, yields a velocity parameter $\xi_0 = 1.6$ km s^{-1}. As the thermal velocity for these atoms in the solar atmosphere is about 1.2 km s^{-1}, this implies a microturbulent velocity of about 1.0 km s^{-1}. By a fit to the damping part of the theoretical curves, $\log a = -1.4$ was obtained. Adopting an average wavelength of 4500 Å for the lines, this yields $\Gamma = 1.7 \times 10^9$ s^{-1}, which is very nearly a factor of 10 larger than the classical damping constant Γ_c. It is clear that the main source of the line broadening must be collisions,

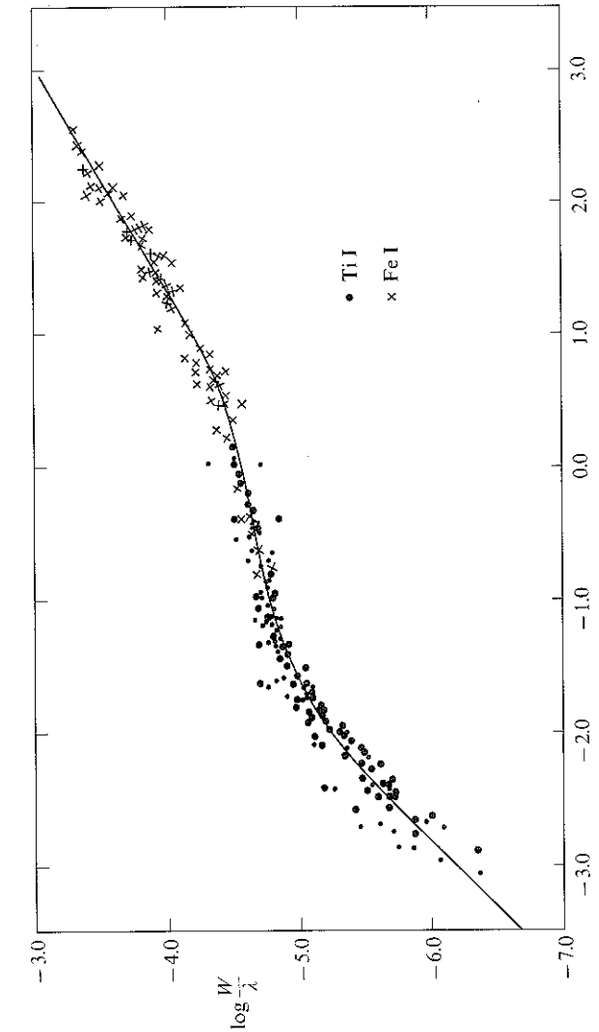

FIGURE 10-3
Empirical curve of growth for solar Fe I and Ti I lines. Abscissa is based on laboratory f-values. From (**686**).

and because hydrogen is mostly un-ionized in the solar photosphere, the most likely process is van der Waals interactions between hydrogen and the radiating atom. An approximate result for Γ_6 was obtained previously [equation (9-71)], namely $\Gamma_6 \approx 8.1\ C_6^{\frac{2}{5}} v^{\frac{3}{5}} N_H$ where $C_6 \approx [13.6/(\chi_{ion} - \chi_i)]^2 \times 10^{-32}$ [equation (9-76)]. For a typical value $(\chi_{ion} - \chi_i) = 6$ eV, $C_6 \approx 5.2 \times 10^{-32}$; typical values for v and N_H are 10^6 cm s^{-1} and 10^{17} cm^{-3}, so $\Gamma_6 \approx 10^9$, which is in basic agreement with the empirical value. The damping part of the empirical curve is found to have a slope (on a logarithmic scale) closer to 0.6 than 0.5. The likely explanation is that lines of widely differing strengths are formed in different layers, so that the parameters describing them (e.g., damping widths) are not identical, as was assumed in constructing the curve of growth.

Many other analyses of the solar spectrum have been made. An improvement over the basic curve of growth method is obtained by using a *saturation function* (**505**) that accounts for the fact that the line cores and wings are actually formed in different layers of the atmosphere. This is conveniently done in the *method of weighting functions*, which is described in detail in (**261**, Chap. 4), and which has been applied in one of the classic abundance analyses of the solar spectrum (**252**).

An important application of the solar curve of growth is in *differential abundance analyses* of stars with respect to the Sun. Excellent examples of this approach for G-type subdwarfs (extremely metal-deficient stars) and for Population II K-giants can be found in (**13**) and (**294**) respectively. An advantage of this approach is that oscillator strengths, which are often poorly known, cancel out of the analysis to first order. Also, because the temperatures of these stars are close to the solar value, one might hope that their atmospheric structures are at least roughly similar to the Sun's. If this is the case, then other effects, such as line-blending, stratification, departures from LTE, etc. might also cancel out to first order.

The fundamental assumption made in a differential analysis is that the *stellar* curve of growth, $\log(Wc/\lambda\xi_0)^*$ versus $\log \beta_0^*$, is identical to the *solar* curve, $\log(Wc/\lambda\xi_0)^\odot$ versus $\log \beta_0^\odot$. In practice we do not know β_0^* or ξ_0^*, and thus we cannot construct the stellar curve directly. What *is* known for each line is $\log(W/\lambda)^*$ and the value of $\log \beta_0^\odot$ for that line from its observed solar equivalent width. From the definition of β_0 [see equation (10-45)], we have

$$\log\left(\frac{\beta_0^\odot}{\beta_0^*}\right) = \log\left(\frac{N_{jk}^\odot}{N_{jk}^*}\right) + \chi_{ijk}(\theta_{exc}^* - \theta_{exc}^\odot) + \log\left(\frac{\xi_0^* k_c^* U^*}{\xi_0^\odot k_c^\odot U^\odot}\right) \quad (10\text{-}50)$$

If we define $[X] \equiv \log(X^\odot/X^*)$ for any quantity X, then equation (10-50) may be rewritten as

$$[\beta_0] = [N_{jk}] + \chi_{ijk}\,\Delta\theta - [\xi_0] - [k_c] - [U] \quad (10\text{-}51)$$

Thus, if we plot $\log(W/\lambda)^*$ for lines of a given ion, say Fe I, versus $\log \beta_0^\odot$ instead of $\log \beta_0^*$, lines of different excitation potentials will scatter around a mean curve if there is a difference in excitation temperature between the star and the sun. Therefore, we plot $\log(W/\lambda)^*$ versus $(\log \beta_0^\odot - \chi_{ijk}\,\Delta\theta)$ and choose $\Delta\theta$ to minimize the scatter. For the G-subdwarfs it is found (**13**) that $\Delta\theta \lesssim 0.05$, while for Population II K-giants $\Delta\theta$ is about 0.25 to 0.35 (**294**). After the effect of $\Delta\theta$ is eliminated, the empirical curve for the star is superimposed upon the solar curve, in which $\log(Wc/\lambda\xi_0)^\odot$ is plotted versus $\log \beta_0^\odot$. The *vertical* shift gives directly $[\xi_0]$ while the *horizontal* shift yields the average value of $\delta \equiv [N_{jk}] - [\xi_0 k_c U]$ for the ion under consideration. For the G-subdwarfs ξ_0 is found to correspond closely to a pure thermal value (no turbulence), while for the K-giants there is appreciable turbulence. As the temperatures of these stars are fairly close to the solar value, the partition-function ratio is usually set to unity, and the value of δ depends primarily upon abundances, and upon differences in ionization and continuous opacity. Because H^- is the main contributor to k_c in the temperature range under consideration, the value of $[k_c]$ will essentially equal $[n_e]$. We can determine $[n_e]$ from an analysis of the ionization equilibrium using information from two stages of ionization of the same element k. If we have observations for two stages of ionization (say "0" and "1"), we may write

$$\delta_{jk} = [N_{jk}] - [\xi_0 k_c], \quad (j = 0, 1) \quad (10\text{-}52)$$

Using these two values of δ we can derive

$$\begin{aligned}\Delta \equiv \delta_{1k} - \delta_{0k} &= [N_{1k}] - [N_{0k}] \\ &= \log(N_{1k}^\odot/N_{0k}^\odot) - \log(N_{1k}^*/N_{0k}^*)\end{aligned} \quad (10\text{-}53)$$

As $\log(N_{1k}^\odot/N_{0k}^\odot)$ is known, the observed value of Δ yields $\log(N_{1k}^*/N_{0k}^*)$. But if we assume that the temperature is specified by θ_{exc}^* (which is known), then Saha's equation gives $\log(N_{1k}^* n_e^*/N_{0k}^*)$; hence we may determine $\log n_e^*$ and $[n_e]$. Estimates of $[n_e]$ can be obtained from several different elements, and a mean value taken. Knowledge of $[n_e]$ allows us to evaluate $[k_c]$ and $\log(N_{jk}^*/N_k^*)$. Hence finally we can calculate the ratio of the abundance of the element in the star to its abundance in the Sun as

$$\begin{aligned}\log(N_k^\odot/N_k^*) &= \log(N_k^\odot/N_{jk}^\odot) + \log(N_{jk}^*/N_k^*) + \log(N_{jk}^\odot/N_{jk}^*) \\ &= \log(N_k^\odot/N_{jk}^\odot) + \log(N_{jk}^*/N_k^*) + \delta + [\xi_0 k_c U]\end{aligned} \quad (10\text{-}54)$$

because all four terms on the righthand side of the second equality are now known.

The results of the two analyses mentioned above lead to the striking conclusion that the abundances of the heavy elements in Population II stars are lower than their abundances in the Sun by a factor of the order

of 10^2! This is a fact of tremendous significance for the construction of a picture of the evolution of the Galaxy, and it implies that the heavy elements in Population I stars such as the Sun are the result of nucleosynthesis in earlier generations of stars. The errors in abundances derived from curve-of-growth studies are typically of the order of a factor-of-two uncertainty; although such errors are serious, and motivate the development of more precise methods, they do not change the qualitative results for Population II stars mentioned above.

10-5 LTE Spectrum Synthesis with Model Atmospheres

The curve of growth approach introduces a large number of simplifying approximations, which seriously limit the accuracy of the results obtained Considerable improvement is achieved if the *physical assumption* of LTE is retained, but now a *model atmosphere* is employed to represent the depth-variation of all physical parameters. Given such a model, one may compute both κ_c and $\chi_l\phi_\nu$ as functions of depth, allowing fully for the variation of the temperature, ionization–excitation equilibrium, Doppler widths, damping parameters, etc. It is then possible to compute the optical depth

$$\tau_\nu(z) = \int_z^{z_{\max}} [\kappa_c(z') + \chi_l(z')\phi_\nu(z')]\, dz' \tag{10-55}$$

and hence the emergent flux in a line at frequency ν

$$F_\nu = 2\int_0^\infty B_\nu[T(\tau_\nu)]E_2(\tau_\nu)\, d\tau_\nu \tag{10-56}$$

by direct numerical quadrature, with as much mathematical accuracy as desired.

The first step in this method is to select the model atmosphere that most closely resembles the stellar atmosphere to be analyzed. This choice is made by comparing observed and computed values of certain key features in the spectrum. Typically the comparison is made for (1) continuum features such as the overall energy distribution, the Balmer jump D_B, or a color such as $(b - y)$; (2) line profiles of the hydrogen lines (which are density-sensitive); (3) ratios of line-strengths for lines of two ionization stages of a given element. For example, in B-stars, the continuum parameters determine mainly the effective temperature T_{eff}; the hydrogen lines determine mainly $\log g$; the ratio of, say, Si III λ4552 to Si II λ4128, 31 is a function of both T_{eff} and $\log g$ (the *ratio* is insensitive to the element abundance). Examples of fits of computed and observed energy distributions are shown in Figures 7-4 through 7-6 and in Figure 7-8. Figure 10-4 shows a fit to the

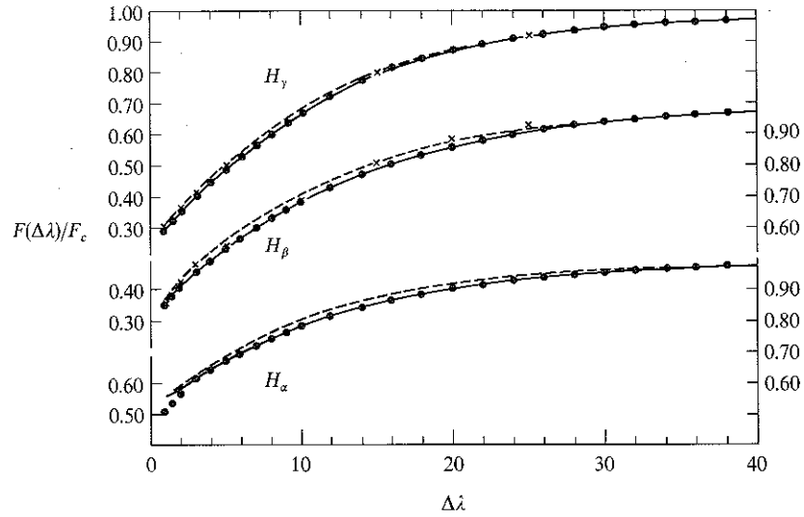

FIGURE 10-4
Hydrogen Balmer-line profiles in Vega. *Solid dots*: observed profiles from (**509**). *Crosses*: observed profiles from (**71**). *Abscissa*: $\Delta\lambda$ in Å. Curves are computed profiles from a model with $T_{\text{eff}} = 9650°$K and $\log g = 4.05$. From (**555**), by permission.

observed hydrogen line profiles in the standard star Vega (**555**). Often a plot is made, in a diagram of T_{eff} versus $\log g$, of the loci of (T_{eff}, g) values for which the computed value of a given parameter [say D_B or $W(H\gamma)$] equals the observed value. The loci for different criteria will intersect, and thus define an optimum value of (T_{eff}, g). Ideally the curves would intersect at a single point; in practice they will intersect within a small area, which introduces an uncertainty into both T_{eff} and g. Examples of this procedure for the normal B-star γ Peg and the He star HD 184927 are given in (**508**) and (**299**), respectively.

Having chosen a model atmosphere, it becomes possible to carry out an abundance analysis. One now replaces the curve of growth by a computation of the equivalent width, of each line under consideration, as a function of element-abundance relative to hydrogen. Knowledge of the observed equivalent width of each line thus leads to an estimate of the element abundance; these estimates may be suitably weighted and averaged over all lines to yield a final abundance estimate.

Although the choice of the model atmosphere used in the abundance analysis is based on continuum and hydrogen-line criteria, there is usually some ambiguity in $(T_{\text{eff}}, \log g)$. Often the results of the abundance analyses for several elements can be used to narrow the range of uncertainty in

(T_{eff}, log g), and to choose a more refined model, much in the same way as atmospheric properties may be inferred from a curve-of-growth analysis. For example, after the analysis is completed, the derived abundances for many lines may be examined to see if they show a correlation with excitation potential; such a correlation would be expected if the temperatures in the atmosphere are incorrectly chosen. The elimination of any such correlations may, therefore, lead to an improved choice of T_{eff} for the model. Similarly, for a given element, identical abundances should be derived from all ionization stages. Differences that are found between ionization stages contain information about errors in the choice of temperatures and pressures, and thus about T_{eff} and log g for the model. Finally, one can examine the derived abundances to see if they show a correlation with equivalent width. If, for example, stronger lines systematically lead to larger abundances, the velocity parameter may have been underestimated, and thus information about turbulent velocities to be included in the model can be derived. (As we shall see in Chapter 11, however, departures from LTE must also be considered when a determination of the velocity parameter from strong lines is attempted.) Naturally, errors in the observed equivalent widths, in f-values, and in other atomic parameters will introduce scatter; hence the procedure outlined above may not be completely unambiguous. Nevertheless, examination of the correlations just mentioned can often lead to a significantly better choice for the model. Finally, if the derived abundances are very different from those used in construction of the model (as they might be for, say, Ap stars), it may be necessary to recompute models with appropriate abundances, re-derive T_{eff} and log g, and perform the analysis over.

An example of the model-atmosphere method of abundance analysis is that for the two bright A-stars Vega and Sirius in (**609**). The results of this work show that Vega has very nearly solar abundances of the heavy elements, while Sirius has abundances a factor of 4 to 10 higher. In many respects Sirius mildly resembles the group of A-stars called *metallic-line A-stars* (Am stars) because of the strength of the metal lines in their spectra. The present literature of abundance analyses (and estimates of atmospheric parameters such as T_{eff} and log g) by both curve-of-growth and model-atmosphere techniques is vast; extensive lists of references to research papers on the subject for both normal and peculiar stars can be found in (**12**; **70**, 57–204; **144**; **450**, 157–237; **493**; **523**; **552**; **560**; **658**). The work thus far described makes use primarily of equivalent widths. For some stars, for which very high-quality spectroscopic data are available, a detailed point-by-point synthesis of the spectrum is possible. An example of such a synthesis for the Sun (**538**) is shown in Figure 10-5. Analyses of this kind can lead to accurate abundances and to the identification of hitherto unnoticed weak blends in the spectrum; applications of spectrum synthesis to the calibration of photometric systems were described in §7-4.

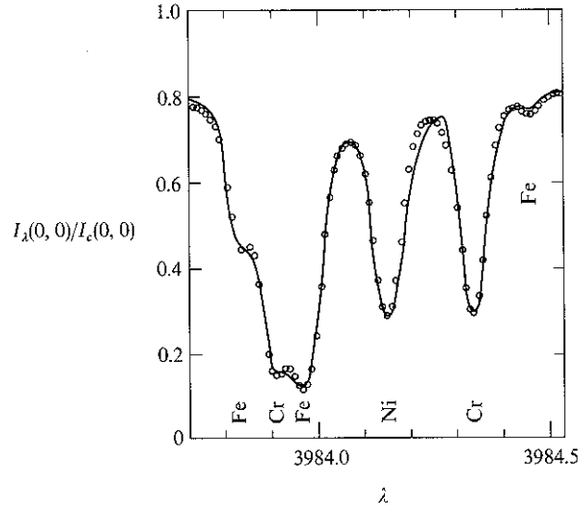

FIGURE 10-5
Spectrum synthesis for solar Fe, Cr, and Ni lines. *Curve*: observed intensity; *open circles*: computed intensity. *Ordinate*: emergent specific intensity at disk-center relative to continuum; *abscissa*: wavelength in Å. From (**538**) by permission.

All of the methods described in this chapter make the assumption of LTE. As we shall see in Chapters 11 and 12, however, this assumption is often a poor description of the physics of line-formation, and can lead to results containing serious systematic errors. Let us now, therefore, turn to the problem of solving the combined equations of radiative transfer and atomic statistical equilibrium fully consistently.

11

Non-LTE Line Transfer: The Two-Level Atom

Much of the progress in understanding the physics of line-formation has come from a study of solutions of the *combined* transfer and statistical equilibrium equations: the so-called non-LTE approach. In this chapter we shall consider some schematic line-formation problems that are simple enough to be solved readily, but which, nevertheless, provide a good description of the physically important processes—and yield considerable insight. It will be obvious that some of the assumptions made are over-simplifications, and are not valid in actual stellar atmospheres, for which elaborate numerical calculations are required to yield accurate results. On the other hand, the real goal is to *understand* the answers, not merely obtain them. This can be done only with a clear grasp of the prototype cases discussed in this chapter, which provide a *conceptual* framework of great usefulness for the interpretation of results from computations with complicated model atoms and detailed model atmospheres.

11-1 Diffusion, Destruction, Escape, and Thermalization

The most important difference between the LTE and non-LTE treatments of line-formation is the way in which the coupling between the gas and radiation

is handled. In the LTE approach, it is *assumed* that the *local* values of two thermodynamic variables (T and N) are sufficient to determine completely the excitation and ionization state of the gas (and hence transfer quantities such as χ_ν, η_ν, and S_ν), independently of the state of the gas at other points. (This remark is strictly true only when the temperature structure is *given*; if it is determined by the condition of radiative equilibrium, then there is a coupling among different points in the atmosphere—see §7-2—but not in the sense we are concerned with here.) It has been emphasized in previous chapters that the excitation and ionization state of the gas is, in fact, strongly influenced by the radiation field, which, in turn, is determined by the state of the gas throughout large volumes of the atmosphere, via the transfer equation. In actuality the two problems of radiative transfer and statistical equilibrium are inextricably coupled, and must be considered *simultaneously*.

Much of the basic physics of line-formation can be understood by consideration of the characteristic lengths for photon diffusion, destruction, and thermalization, which are related to photon escape and destruction probabilities. For simplicity we shall suppose that the atom consists of a ground state (denoted l) and a single excited state (denoted u). The distance over which a line photon will, on the average, move in the atmosphere between successive interactions with the material (absorptions) can be represented by the average *mean-free-path* l. The mean-free-path for a photon emitted at a given frequency is the distance interval whose optical thickness at that frequency is about unity—i.e., $l_\nu \approx 1/\chi_\nu = (\chi_{lu}\phi_\nu + \chi_c)^{-1}$ where χ_{lu} and χ_c are the line and overlapping continuum opacities, and ϕ_ν is the line profile. For complete redistribution, the probability of emission at frequency ν is ϕ_ν, so $l = \langle l_\nu \rangle = \int l_\nu \phi_\nu \, d\nu$. It is obvious that photons emitted in the core travel relatively small distances, while those in the wings can travel much greater distances, up to a distance corresponding to unit optical thickness in the continuum.

When a photon is absorbed in the line and excites an atom to the upper state, it is *usually* re-emitted in a radiative de-excitation and travels another mean-free-path. This process may occur *again and again* before the photon is ultimately *destroyed*, and has its energy deposited into the thermal pool, either by a collisional de-excitation or by an absorption in the overlapping continuum. Thus there exists a characteristic length L, the *destruction length*, over which a photon may travel before it is destroyed. The destruction length has a more basic physical significance than the mean-free-path, for it measures the distance over which a photon emitted at a given point retains its identity and hence can "communicate" information about conditions at that point to another. Thus L determines an *interaction region*: the volume containing those points that can influence one another via photon exchange.

The relative sizes of L and l depend upon the photon *destruction probability* P_d, which gives the average probability that the photon is destroyed

when it next interacts with the material. The probability for photon destruction by collisions following absorption in the line is $C_{ul}/(A_{ul} + C_{ul})$ where C_{ul} is the rate of collisional de-excitation and A_{ul} is the spontaneous emission rate; now $C_{ul} \propto n_e$, so that it is clear that the contribution by collisions to the total destruction probability becomes large (i.e., approaches unity) in the deeper, denser regions of a stellar atmosphere, and can be quite small in the uppermost layers. If we assume that all photons absorbed in the continuum are destroyed thermally, the contribution of continuum processes to P_d is the average of $\chi_c/(\chi_{lu}\phi_\nu + \chi_c)$ over the line profile. The continuum sets an upper bound on both l and L because a photon cannot travel more than a unit optical depth in the continuum before it is absorbed—and when it is, it is also destroyed.

At great depths in the atmosphere, P_d approaches unity because of large densities (and hence collision rates) and because of the strength of the overlapping continuum. Then $L \to l$, for the photon is almost surely destroyed when it is next absorbed. Assuming that the continuum is already optically thick, it is clear that the radiation field becomes very strongly coupled to local conditions and *thermalizes* to its local thermodynamic equilibrium value. In contrast, when the destruction probability is very small, $L \gg l$, and the interaction region may become *enormous* compared to the volume over which a photon can diffuse in a single flight. In this case, the radiation field is dominated by *nonlocal* influences, and represents the result of physical conditions that may be quite inhomogeneous. For example, within the volume there may be large variations of the kinetic temperature that imply strong variations of thermodynamic properties and of the Planck function. The radiation field may then depart *markedly* from its local equilibrium value, and this departure will extend throughout the entire interaction region.

The importance of these notions becomes manifest when we consider a sequence of test points approaching the surface of the atmosphere. At the deepest points we obtain equilibrium. But as we approach the surface, the distance L grows, and eventually the interaction volume contains the boundary itself (more precisely, extends above the depth at which $\tau = 1$ anywhere in the line). Then a new phenomenon enters the picture: photons *escape* from the medium into space. It follows that the radiation field, at test points whose interaction regions extend into space beyond the surface, must be depressed *below* the value it would have had if there were no boundary, for no radiation is emitted in the region void of material, and photon escapes are therefore uncompensated. This effect leads to a deficiency in the radiation throughout an entire interaction volume, and hence extends to depths at least of the order of L. But of course the radiation field at these points influences that at points that lie yet another destruction length L deeper, and therefore there is a "compounding" of the effect, which leads to

a departure of the radiation field from its equilibrium value at positions far below the surface.

The depth at which the radiation field (or source function) ultimately approaches closely its equilibrium value is the *thermalization depth* Λ; this concept, introduced by Jefferies (**333**), has proven to be extremely fruitful. To obtain a quantitative estimate of Λ we define the *escape probability* $P_e(\tau)$ as the probability, averaged over the line, that a photon emitted at line optical depth τ escapes from the medium before being absorbed along its path of flight. The mechanism competing with photon escape is photon destruction; the former induces a departure from equilibrium while the latter leads to coupling to the local thermal pool. We therefore compare P_e with P_d. Deep within the atmosphere where $P_e(\tau) \ll P_d$, photons are surely thermalized before they can escape, hence $S \to B$; at the surface where $P_e(\tau) \gg P_d$, photons escape freely before thermalization, hence S will depart from B. It is therefore reasonable, on physical grounds, to identify the thermalization depth Λ as that point at which $P_e(\Lambda) = P_d$. Defined in this way, Λ is essentially the greatest depth from which photons have a significant chance to escape before being destroyed.

Now the escape probability (and hence Λ) depends sensitively (a) upon the nature of photon redistribution over the line profile upon emission, and (b) upon the amount of background continuum absorption. The probability for photon absorption is highest at line center. If photons are emitted *coherently*, then those absorbed at line center will be re-emitted there, and hence will tend to be trapped by the large line-core optical depth up to the very uppermost layers, where the line finally becomes optically thin; this will tend to prevent serious departures of the line-core radiation field (which contributes heavily to the total photoexcitation rate) from its equilibrium value up to the shallowest layers of the atmosphere. In contrast, if photons are *completely redistributed* over the line profile, then there is a significant chance that after a number of scatterings a photon absorbed at line-center will be emitted in the line-wing where the opacity is low, and the probability of escape is high. Photons that would have been trapped, if emitted coherently, now freely escape from the atmosphere, depressing the line-core intensity (and photoexcitation rate) in much deeper layers in the atmosphere. The radiation field in the line as a whole now responds to the fact that the boundary lies within a mean-free-path at some frequencies, even if not at others. The role of the continuum is obviously important in the case of complete redistribution, for it sets a lower bound on the total opacity, and hence an upper bound on the depth from which escapes become possible at *any* frequency, no matter how far out in the line-wing. It is thus clear that we should expect the thermalization depth to be much larger, and the surface departures from equilibrium larger, when the scattering process is noncoherent instead of coherent; further, we expect the magnitude of these

effects to be larger, the more highly developed the line-wings are relative to the core. We shall find these expectations are fulfilled by the detailed analysis given below.

It has been possible to extract practically all of the physical flavor of the problem using the qualitative arguments given above. Let us now turn to a mathematical discussion to obtain quantitative results, and to extend the analysis to cases where such heuristic discussion becomes less effective.

11-2 The Two-Level Atom without Continuum

THE SOURCE FUNCTION

Consider a schematic atomic model consisting of only *two levels*, l and u, between which radiative and collisional transitions can occur. This model is obviously very incomplete, but it nevertheless provides a fairly good description of the real situation for some lines. In particular, *resonance lines* arising from the ground state are well described by this model when the coupling of the lower and upper levels to the continuum (and of the upper level to other levels) is weak. For the present, assume that the *only* sources of opacity and emission at the line frequency are from the line itself; then the transfer equation is

$$\mu(dI_\nu/dz) = [-n_l B_{lu} I_\nu + n_u(A_{ul} + B_{ul} I_\nu)]\phi_\nu(h\nu/4\pi) \tag{11-1}$$

Here we have assumed *complete redistribution* so that the absorption and emission profiles are identical. Define the optical depth scale in terms of the frequency-integrated line opacity (which characterizes the average opacity in the line as a whole), $d\tau \equiv -\chi_{lu}\, dz$ where

$$\chi_{lu} = (n_l B_{lu} - n_u B_{ul})(h\nu/4\pi) = (\pi e^2/mc) f_{lu}[n_l - (g_l/g_u)n_u] \tag{11-2}$$

Then

$$\mu(dI_\nu/d\tau) = \phi_\nu(I_\nu - S_l) \tag{11-3}$$

where

$$S_l = n_u A_{ul}/(n_l B_{lu} - n_u B_{ul}) = (2h\nu^3/c^2)[(n_l g_u/n_u g_l) - 1]^{-1} \tag{11-4}$$

the second equality following from application of the Einstein relations [cf. equations (4-8) and (4-9)]. As the factor ν^3 varies only negligibly over the sharply peaked profile ϕ_ν, S_l is called the *frequency-independent* source function; when the emission profile differs from the absorption profile, S_l becomes explicitly frequency-dependent (see §2-1 and §13-4). Equation (11-4) is an *implicit* form for the source function because the level-populations depend upon the radiation field; this dependence can be displayed *explicitly* by incorporating information from the equations of statistical equilibrium that determine n_l and n_u.

The statistical equilibrium equation for level l is

$$n_l\left(B_{lu}\int \phi_\nu J_\nu\, d\nu + C_{lu}\right) = n_u\left(A_{ul} + B_{ul}\int \phi_\nu J_\nu\, d\nu + C_{ul}\right) \tag{11-5}$$

This equation can be used to write an expression for (n_l/n_u). Upon substituting this expression into equation (11-4), making use of the detailed-balance result that $C_{lu} = (n_u/n_l)^* C_{ul}$, and using the Einstein relations, we obtain quite directly

$$S_l = \left[\int \phi_\nu J_\nu\, d\nu + \varepsilon' B_\nu\right] \Big/ (1 + \varepsilon') \equiv (1 - \varepsilon)\bar{J}_\nu + \varepsilon B_\nu \tag{11-6}$$

where

$$\varepsilon' \equiv C_{ul}(1 - e^{-h\nu/kT})/A_{ul} \tag{11-7}$$

and

$$\varepsilon \equiv \varepsilon'/(1 + \varepsilon') \tag{11-8}$$

Exercise 11-1: Derive equation (11-6) in detail.

Each of the terms in equation (11-6) admits a straightforward physical interpretation. The source function contains a *noncoherent scattering term* $\bar{J}$ and a *thermal source* term $\varepsilon' B_\nu$. The thermal source term represents photons that are *created* by collisional excitation, followed by radiative de-excitation. The term ε' in the denominator is a *sink term* that represents those photons that are *destroyed* by a collisional de-excitation following a photoexcitation. These two terms describe completely the coupling of the radiation field to the local thermal state of the gas. The scattering term may be viewed as a *reservoir term* that represents the end result of the *cumulative* contributions of the source and sink terms *over the entire interaction region.*

It is clear that, if densities are made sufficiently large, then the collision rate C_{ul} may eventually exceed A_{ul}, so that ε' becomes $\gg 1$; then $S_l \to B_\nu(T)$, and LTE is recovered. However, in virtually all situations of astrophysical interest, $\varepsilon' \ll 1$ in regions of line-formation, and, in general, the source function cannot be expected to have a value close to the Planck function. This state of affairs was partially recognized in the classical theory by the division of lines into the categories of "absorption" and "scattering" lines; this division, however, was largely ad hoc, and the thermal coupling parameter had to be guessed from heuristic arguments. In the present analysis the coupling parameter follows directly, and uniquely, from the statistical equilibrium equation. Further, in the classical treatments of "pure" scattering lines, it was sometimes incorrectly argued that the small thermal terms could be discarded. The important point to bear in mind is that, even if the thermal term $\varepsilon' B$ is small compared to the scattering term $\bar{J}$ locally, when integrated over the entire interaction region it accumulates to a value of importance;

moreover, at depths greater than the thermalization depth the intensity must ultimately couple to the thermal pool. Indeed, if the thermal terms are discarded, the transfer equation becomes *homogeneous* in the radiation field and the scale of the solution is *unknown*; this scale is, in fact, *fixed by the* (small) *thermal terms* at the point of thermalization.

SOLUTION OF THE TRANSFER EQUATION

Having obtained an expression for the source function, let us now consider the solution of the transfer equation. It is convenient to work with the dimensionless frequency-variable x, measured from line center in units of Doppler widths (or damping widths for Lorentz profiles). In terms of this variable we shall write Doppler profiles as

$$\phi(x) = \pi^{-\frac{1}{2}} e^{-x^2} \tag{11-9a}$$

Voigt profiles as

$$\phi(x) = a\pi^{-\frac{3}{2}} \int_{-\infty}^{\infty} e^{-y^2} [(x - y)^2 + a^2]^{-1}\, dy \tag{11-9b}$$

and Lorentz profiles as

$$\phi(x) = (1/\pi)/(1 + x^2) \tag{11-9c}$$

which are all normalized such that

$$\int_{-\infty}^{\infty} \phi(x)\, dx = 1 \tag{11-10}$$

We will absorb a factor of $\Delta\nu_D$ [or $(\Gamma/4\pi)$ for a Lorentz profile] into the definition of χ and write $d\tau \equiv -\chi_l(z)\, dz$, where now $\chi_l \equiv \chi_{lu}/\Delta\nu_D$, χ_{lu} being given by equation (11-2). Then

$$S_l(\tau) = [1 - \varepsilon(\tau)] \int_{-\infty}^{\infty} \phi(\tau, x) J(\tau, x)\, dx + \varepsilon(\tau) B(\tau) \tag{11-11}$$

and the transfer equation becomes

$$\mu[dI(\tau, x)/d\tau] = \phi(\tau, x)[I(\tau, x) - S_l(\tau)] \tag{11-12}$$

We now introduce discrete angle and frequency meshes $\{\mu_m\}$ and $\{\nu_n\}$, and replace *integrals* with *quadrature sums*:

$$\int_{-\infty}^{\infty} \phi(x) f(x)\, dx \to \sum_{n=-N}^{N} a_n f(x_n) \tag{11-13a}$$

and

$$\int_{-1}^{1} f(\mu)\, d\mu \to \sum_{m=-M}^{M} b_m f(\mu_m) \tag{11-13b}$$

where the points are chosen symmetrically—i.e., $x_{-n} = -x_n$ and $\mu_{-m} = -\mu_m$ (also $a_{-n} = a_n$ and $b_{-m} = b_m$). Then the source function can be written in the same discretized form as was used in Chapter 6 [cf. equation (6-24a)],

$$S_d = (1 - \varepsilon_d) \sum_n a_n \sum_m b_m u_{dmn} + \varepsilon_d B_d \tag{11-14}$$

and the transfer equation can be reduced to the standard *second-order differential equation* form:

$$\mu_m^2 (d^2 u_{mn}/d\tau^2) = \phi_n^2 (u_{mn} - S) \tag{11-15}$$

where $u_\nu \equiv \frac{1}{2}[I_\nu(+\mu) + I_\nu(-\mu)]$. Equation (11-15) can be discretized on a mesh $\{\tau_d\}$, and solved numerically using either the Feautrier or Rybicki schemes discussed in §6-3. All of the relevant parameters (ε, B, and ϕ) may be depth-dependent without causing any difficulty in the calculation.

Alternatively, the transfer equation can be written in *integral equation* form. The formal solution of equation (11-12) is

$$J(\tau, x) = \frac{1}{2} \int_0^\infty S_l(t) E_1 \left| \int_\tau^t \phi(t', x)\, dt' \right| \phi(t, x)\, dt \tag{11-16}$$

For a depth-dependent profile, the argument of the E_1 function will depend upon both τ and t; this greatly complicates the analysis, so we shall consider only the case of a depth-independent profile, in which case only the *displacement* $(t - \tau)$ enters:

$$J(\tau, x) = \frac{1}{2} \phi(x) \int_0^\infty S_l(t) E_1 \left| (t - \tau)\phi(x) \right| dt \tag{11-17}$$

Substituting this expression for J into equation (11-11) yields an integral equation for S_l,

$$S_l(\tau) = [1 - \varepsilon(\tau)] \int_0^\infty S_l(t) K_1 \left| t - \tau \right| dt + \varepsilon(\tau) B(\tau) \tag{11-18}$$

where the *kernel function* K_1 is given by

$$K_1(s) \equiv \frac{1}{2} \int_{-\infty}^{\infty} E_1(\phi_x s) \phi_x^2\, dx = \int_0^\infty E_1(\phi_x s) \phi_x^2\, dx \tag{11-19}$$

Exercise 11-2: Fill in the missing steps in the derivation of equation (11-18).
Exercise 11-3: (a) Show that the *normalization* of $K_1(s)$ is $\int_0^\infty K_1(s)\,ds = \frac{1}{2}$. (b) Using the result in (a), show that, for $\tau \to \infty$, where $S_l(\tau)$ must be a slowly-varying function and hence can be removed from the integral, $S_l(\tau) \to B$.

Numerical methods for solving equations of the form of (11-18) are discussed in detail in (**52**) and (**18**, Chap. 8); in most respects these resemble the Rybicki method of solving the differential equations, though Rybicki's method is simpler to apply when the profile is depth-dependent. In essence the solution is obtained by introducing a functional representation of $S(\tau)$ on a discrete grid $\{\tau_d\}$; these functions are integrated *analytically* against the kernel to construct a final matrix system of the form

$$\mathbf{S} = \mathbf{KS} + \mathbf{L} \tag{11-20}$$

where $\mathbf{S} = (S_1, \ldots, S_d, \ldots, S_D)^T$ represents the depth-variation of the source-function. The system is then solved by standard numerical techniques.

One of the advantages of the integral-equation formulation is that it displays explicitly the intimate dependence of the source function upon the mathematical behavior of the kernel function, and considerable insight can be gained by analytical study of the kernel. In particular, the *asymptotic form* of $K_1(s)$, for $s \gg 1$, shows that a characteristic feature of line-formation problems with noncoherent scattering is an extremely *long-range interaction* of one part of the atmosphere with another. Recall that for *coherent scattering* the appropriate kernel function in the formal solution is $E_1\,|t - \tau|$ [cf. equation (2-57) and Exercise 2-10]. This kernel decays very rapidly, falling off as $\exp(-|t - \tau|)/|t - \tau|$, thereby limiting severely the range of depth-points that are directly coupled together in the scattering process. In contrast, the asymptotic behavior of $K_1(s)$ is found to be (**53**, Appendix I) $K_1(s) \sim 1/\{4s^2[\ln(s/\pi^{\frac{1}{2}})]^{\frac{1}{2}}\}$ for Doppler profiles, $K_1(s) \sim a^{\frac{1}{2}}/(6s^{\frac{3}{2}})$ for Voigt profiles, and $K_1(s) \sim 1/(6s^2)$ for Lorentz profiles. Clearly the range of these kernel functions is very large compared to that for coherent scattering, and this implies that the radiation fields at widely separated points in the medium become mutually interdependent.

The physical reason for the long range of the noncoherent kernel functions is, of course, the redistribution of photons into the transparent line-wings. In the coherent case, any photon absorbed in the core will be re-emitted there, and hence always encounters high opacity and remains trapped. In the noncoherent case it has a probability of being emitted in the wings where, because of the lower opacity, it can travel over a large geometrical distance, and a correspondingly large integrated line optical depth, before it is again absorbed.

THE THERMALIZATION DEPTH

Let us now consider the behavior of the source function in an atmosphere in which ε and B are constant with depth. Naturally this is a very schematic case, but it provides useful insight. One of the most important characteristics of the problem to be determined is the *thermalization depth* at which the line source function approaches closely the Planck function B. Expressions for the thermalization depth can be derived by a wide variety of methods including (a) analysis of the asymptotic behavior of the solution of the integral equation (11-18) (**53**); (b) calculation of the distribution of the distances photons can travel from their points of creation to their places of destruction (**545**; **322**); and (c) calculations of the probability distributions for photon escape from a given point of origin in the medium (**214**; **215**; **216**). One of the simplest and most appealing derivations can be made on the basis of the physical arguments given in §11-1: that the thermalization depth Λ must be near that point in the atmosphere where the probability of *photon escape*, $P_e(\Lambda)$, following a given scattering event is equal to P_d, the probability of *photon destruction* (**28**).

For the strict two-level atom with no overlapping continuum opacity, the probability of destruction per scattering event is just

$$P_d = C_{ul}/(A_{ul} + C_{ul}) = \varepsilon \tag{11-21}$$

where ε is as defined by equations (11-7) and (11-8) (ignoring stimulated emission for simplicity). The escape probability can be calculated by summing $\exp(-\tau\phi_x/\mu)$, the probability of escape at frequency x along a ray at angle $\cos^{-1}\mu$ to the normal, over all angles and frequencies, with weight ϕ_x (the probability for emission at frequency x within the profile):

$$P_e(\tau) = \frac{1}{2}\int_{-\infty}^{\infty} dx\,\phi_x \int_0^1 e^{-\tau\phi_x/\mu}\,d\mu = \frac{1}{2}\int_{-\infty}^{\infty} dx\,\phi_x \int_1^\infty e^{-(\tau\phi_x)v} v^{-2}\,dv$$

$$= \frac{1}{2}\int_{-\infty}^{\infty} E_2(\tau\phi_x)\phi_x\,dx \tag{11-22}$$

For $\tau \gg 1$, the exponential integral is approximately zero for $|x| \leqslant x_1$, where $\phi(x_1)\tau \equiv 1$, and unity for $|x| > x_1$. Physically this states that photons are trapped in the core, where $\tau\phi_x > 1$, but in the wing where $\tau\phi_x \leqslant 1$, they escape freely. We may thus approximate $P_e(\tau)$ as

$$P_e(\tau) \approx \int_{x_1}^{\infty} \phi(x)\,dx \tag{11-23}$$

The above integral is easy to evaluate for $x_1 \gg 1$ (which it will be when $\tau \gg 1$). By direct substitution of equations (11-9) we find

$$P_e(\tau) = \frac{1}{2}\operatorname{erfc}(x_1) \approx e^{-x_1^2}/(2\pi^{\frac{1}{2}}x_1) \qquad \text{(Doppler)} \qquad (11\text{-}24a)$$

$$P_e(\tau) \approx a/(\pi x_1) \qquad \text{(Voigt)} \qquad (11\text{-}24b)$$

$$P_e(\tau) \approx 1/(\pi x_1) \qquad \text{(Lorentz)} \qquad (11\text{-}24c)$$

In deriving equation (11-24b), use was made of the asymptotic form of the Voigt function [cf. equation (9-45)]. Now the condition $\phi(x_1)\tau \equiv 1$ implies

$$x_1 = [\ln(\tau/\pi^{\frac{1}{2}})]^{\frac{1}{2}} \qquad \text{(Doppler)} \qquad (11\text{-}25a)$$

$$x_1 = (a\tau/\pi)^{\frac{1}{2}} \qquad \text{(Voigt)} \qquad (11\text{-}25b)$$

$$x_1 = (\tau/\pi)^{\frac{1}{2}} \qquad \text{(Lorentz)} \qquad (11\text{-}25c)$$

Hence, substituting equations (11-25) into (11-24), setting $P_e(\Lambda) = P_d = \varepsilon$, and solving for Λ we find

$$\Lambda \approx c/\varepsilon \qquad \text{(Doppler)} \qquad (11\text{-}26a)$$

$$\Lambda \approx a/\varepsilon^2 \qquad \text{(Voigt)} \qquad (11\text{-}26b)$$

$$\Lambda \approx 1/\varepsilon^2 \qquad \text{(Lorentz)} \qquad (11\text{-}26c)$$

where c is a number of order unity that depends implicitly upon ε; other factors of order unity have been suppressed.

Exercise 11-4: Show that for a profile with asymptotic form $\phi(x) \sim cx^{-\alpha}$, with $\alpha > 1$, the thermalization depth varies as $\Lambda \sim \varepsilon^{-\beta}$, where $\beta = \alpha/(\alpha - 1)$.

The striking feature of the results given by equations (11-26) is that the thermalization depth for a line with small ε is *enormous*. Recall that for coherent scattering the thermalization depth $\sim\varepsilon^{-1}$ (cf. §6-1 and 10-2). It is clear that the effects of noncoherence, with their attendant increase of photon diffusion in the line wings, greatly increase the depth in the atmosphere over which the source function can depart from the local Planck function. In fact, intercomparison of equations (11-26a) through (11-26c) shows immediately that an increase in the relative importance of the line-wing within the profile leads directly to an increase in the thermalization depth. It is worth stressing, however, that the above results do explicitly depend upon the assumptions of (a) no background opacity and (b) complete redistribution. We shall see later in this section that the effect of a background opacity can greatly reduce Λ.

In Chapter 13 we shall find that in the case of resonance lines, the scattering in strong *radiation-damping* wings is nearly *coherent*, and Λ is better approximated by equation (11-26a) than (11-26b).

From a mathematical point of view, equations (11-26) show that any *iterative* solution of the transfer equation starting from LTE is *futile* for small ε (much more so than the cases already discussed in §6-1!), and only *direct* solutions of the transfer equation will yield correct values.

BOUNDARY-VALUE AND DEPTH-VARIATION OF THE SOURCE FUNCTION

Having shown above that the source function may be decoupled from the Planck function over great depths in the atmosphere, let us now inquire what value the source function attains near the boundary of a semi-infinite atmosphere. If ε and B are taken to be constant with depth, the transfer equation (11-12) is easily solved by the method of discrete ordinates (**53**; **476**). Writing $\lambda_n \equiv 1/\phi(x_n)$ we have the coupled system of equations

$$\lambda_n\mu_m(dI_{mn}/d\tau) = I_{mn} - \frac{1}{2}(1-\varepsilon)\sum_{n'=-N}^{N} a_{n'}\sum_{m'=-M}^{M} b_{m'}I_{m'n'} - \varepsilon B \qquad (11\text{-}27)$$

Considering first the homogeneous equation ($B \equiv 0$) and seeking a solution of the form $I_{mn} = g_{mn}\exp(-k\tau)$ one finds that

$$I_{mn} = c(1 + k\lambda_n\mu_m)^{-1}e^{-k\tau} \qquad (11\text{-}28)$$

where the constants k are roots of the characteristic equation

$$2(1-\varepsilon)\sum_{n=1}^{N} a_n \sum_{m=1}^{M} b_m(1 - k^2\lambda_n^2\mu_m^2)^{-1} = 1 \qquad (11\text{-}29)$$

Exercise 11-5: (a) Verify equation (11-28). (b) Show that equation (11-29) is the appropriate characteristic equation for the system (11-27).

For $\varepsilon = 0$, the root $k^2 = 0$ is a solution. For $\varepsilon > 0$, k^2 must be greater than zero. As in our earlier use of the discrete-ordinate method (§3-3, §7-4) we may delimit the roots by the poles at $k^2 = 1/\lambda_n^2\mu_m^2$, and if we label these poles in the order of *decreasing* values of $(\lambda_n\mu_m)$ we may write

$$0 < k_1^2 < (\lambda\mu)_1^{-2} < k_2^2 < \cdots < k_{MN}^2 < (\lambda_N\mu_M)^{-2} \qquad (11\text{-}30)$$

which shows that the roots may be determined rapidly by a systematic search on appropriate finite intervals. As can be seen, it is convenient to label the roots with a subscript α, $1 \leqslant \alpha \leqslant MN$. By an analysis of the characteristic

equation similar to that used in deriving equation (3-65) one may readily show that

$$\prod_{\alpha=1}^{MN} (k_\alpha \lambda_n \mu_m)^2 = \varepsilon \tag{11-31}$$

We shall use this result below. For B = constant, a particular solution of equation (11-27) is $I_{mn} = B$. Thus the general solution must have the form

$$I_{mn} = B\left[1 + \sum_{\alpha=1}^{MN} L_\alpha e^{-k_\alpha \tau}(1 + k_\alpha \lambda_n \mu_m)^{-1}\right] \tag{11-32}$$

where the ascending exponentials have been discarded to keep the solution bounded at infinity. The surface boundary condition is, as usual, $I_{mn}(-\mu_m) = 0$. If we define

$$\mathscr{S}(x) \equiv 1 + \sum_{\alpha=1}^{MN} L_\alpha (1 - k_\alpha x)^{-1} \tag{11-33}$$

then the surface boundary condition may be expressed as $\mathscr{S}(\lambda_n \mu_m) = 0$ for $(m = 1, \ldots, M)$ and $(n = 1, \ldots, N)$. This set of linear equations may be solved numerically for the L_α's. Substituting equation (11-32) into the expression for S_l on the right-hand side of equation (11-27), and making use of the characteristic equation (11-29) one finds

$$S_l(\tau) = B\left(1 + \sum_{\alpha=1}^{MN} L_\alpha e^{-k_\alpha \tau}\right) \tag{11-34}$$

In particular, at the surface,

$$S_l(0) = B\left(1 + \sum_{\alpha=1}^{MN} L_\alpha\right) = B\mathscr{S}(0) \tag{11-35}$$

Now, by an analysis similar to that used to derive equations (3-70) and (7-104) one may show that

$$\mathscr{S}(x) = \left[\prod_{\alpha=1}^{MN} k_\alpha \prod_{n=1}^{N} \prod_{m=1}^{M} (\lambda_n \mu_m - x)\right] \Big/ \prod_{\alpha=1}^{MN} (1 - k_\alpha x) \tag{11-36}$$

from which it follows, using equation (11-31) that

$$\mathscr{S}(0) = \prod_{\alpha=1}^{MN} (k_\alpha \lambda_n \mu_m) = \varepsilon^{\frac{1}{2}} \tag{11-37}$$

Exercise 11-6: (a) Derive equation (11-31). (b) Derive equation (11-34). (c) Derive equation (11-36).

By substitution of equation (11-37) into (11-35), we obtain finally the extremely important result

$$S_l(0) = \varepsilon^{\frac{1}{2}} B \tag{11-38}$$

This result is independent of the order of the quadrature sum and the form of the profile, and hence is general. The basic physical content of the results obtained thus far is contained in equations (11-38) and (11-26). It may be summarized in words as follows: when the coupling of the line-emission process to the termal pool is weak ($\varepsilon \ll 1$), the source function can depart *drastically* from the thermal values, and this departure can extend to *great depth* in the atmosphere. Note that the surface value of the source function exceeds the local creation rate by a factor of $1/\varepsilon^{\frac{1}{2}}$. This implies that $S_l(0)$ is controlled primarily by photons fed in from the line wings. These photons originated deep in the atmosphere where $S_l(\tau) \gg S_l(0)$; it is therefore clear that *the surface value of S_l has little to do with the local thermal source term but is dominated by nonlocal effects.*

Numerical solutions (**53**) for the full depth-variation of $S_l(\tau)$ are shown in Figures 11-1 and 11-2. The cases in Figure 11-1 are for a semi-infinite atmosphere, constant Planck function ($B = 1$), a Doppler profile ($a = 0$), and various values of ε. It can be seen that in each case $S_l(0) = \varepsilon^{\frac{1}{2}}$, and that $S_l \to B$ at $\tau \sim 1/\varepsilon$. The results shown in Figure 11-2 are for lines with $\varepsilon = 10^{-4}$, and for Voigt profiles ranging from a pure Doppler profile to a pure

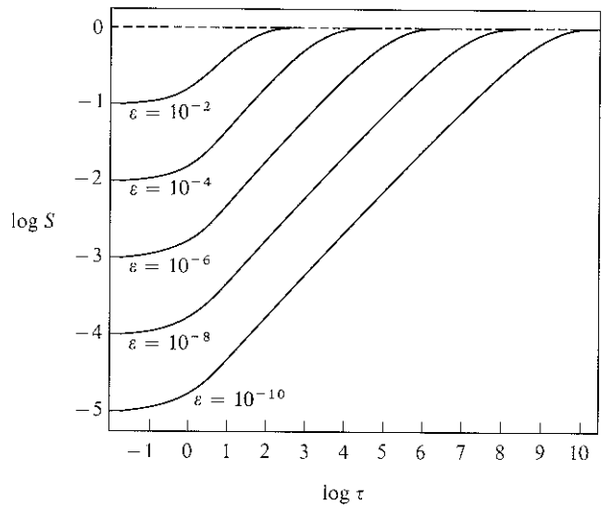

FIGURE 11-1
Line source functions in a semi-infinite atmosphere with $B = 1$, for a line with a pure Doppler profile ($a = 0$), and various values of ε. From (**53**), by permission.

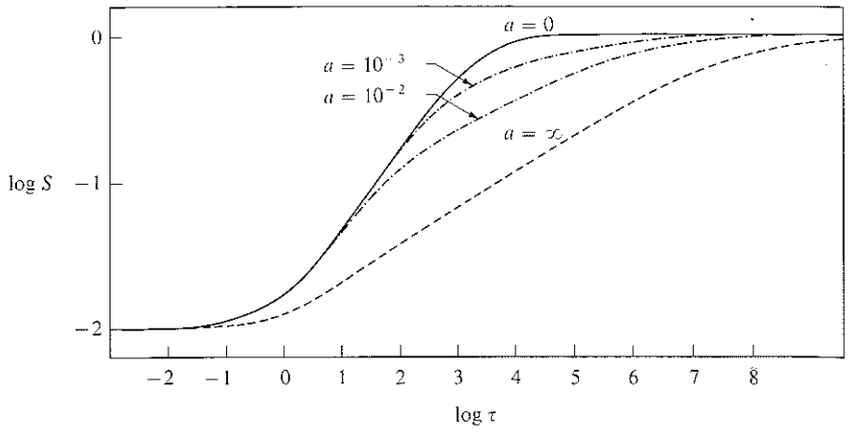

FIGURE 11-2
Line source functions in a semi-infinite atmosphere with $B = 1$, for lines with $\varepsilon = 10^{-4}$, and Voigt profiles ranging from a pure Doppler profile ($a = 0$) to a pure Lorentz profile ($a = \infty$). From (**53**), by permission.

Lorentz profile ($a = \infty$). The increase in thermalization depth from ε^{-1} to ε^{-2} is shown plainly. The source functions in both Figures 11-1 and 11-2 will yield *absorption lines* with dark cores; in contrast, the LTE solution, with $S_l \equiv B$, would yield *no line* whatever. This difference can be attributed to the effects of scattering, and indeed the classical theory would also have predicted a line of the same central depth for the same ε. However there remain two important differences. (a) Both the upper- and lower-state occupation numbers differ from their LTE values; in many earlier "scattering" calculations LTE populations were (incorrectly) assumed. (b) The dark portion of the non-LTE profiles, where *noncoherent* scattering is assumed, are wider. The Eddington–Barbier relation implies that $I_\nu(0) \sim S_l(\tau_\nu = 1)$, hence the lines will be dark for $|x| \lesssim x_1$ where $\Lambda\phi(x_1) = 1$. Because Λ is so much larger for noncoherent scattering (ε^{-1} to ε^{-2} instead of $\varepsilon^{-\frac{1}{2}}$), the corresponding values of x_1 will also be larger; we shall return to this point in §11-5. Finally, it must be remarked that the solutions obtained here are *fully consistent* solutions of both the transfer and statistical equilibrium equations; any approach that falls short of this level of consistency is unsatisfactory.

FINITE SLABS

The *finite slab* atmosphere of total thickness T is a case of astrophysical importance; it can be used to represent nebulae, or limited zones in an

atmosphere to which a particular ion is confined owing to changes in the ionization balance (e.g., chromospheric lines of, say, He II are limited to layers bounded above by the corona and below by the photosphere). In finite atmospheres two physically distinct behaviors are found, depending on whether the atmosphere is *effectively thick or effectively thin*. If $T \gg \Lambda$, then photons from the slab center will not escape before they thermalize; in this case $S_l(0)$ will attain its semi-infinite atmosphere value for the corresponding value of ε, and will approach B at depths $\gtrsim \Lambda$ from the surfaces. If, however, $T \ll \Lambda$, then the solution never thermalizes, and $S_l(\tau)$ becomes proportional to the local creation rate—i.e., $S_l(\tau) = \varepsilon B f(\tau)$ where $f(\tau)$ is independent of ε for a given T.

An estimate of S_l at slab-center can be obtained as follows [see also (**28**)]. The ratio of the *total* number of emissions along a column through the slab to those *thermally created* must be equal to the mean number of times, $\langle N \rangle$, a photon is scattered before it escapes or is destroyed, i.e.,

$$\langle N \rangle = \int_{-\infty}^{\infty} d\nu(4\pi h\nu\phi_\nu) \int_0^T S_l(\tau)\, d\tau \Big/ \int_{-\infty}^{\infty} d\nu(4\pi h\nu\phi_\nu) \int_0^T \varepsilon(\tau)B(\tau)\, d\tau$$
$$= \int_0^T S_l(\tau)\, d\tau \Big/ \int_0^T \varepsilon(\tau)B(\tau)\, d\tau \qquad (11\text{-}39)$$

Here we have used the definition $\eta_\nu = \chi_\nu S_\nu$, and noted that S_l is frequency-independent. For a finite slab the dominant photon loss-mechanism is escape, hence $\langle N \rangle \sim [P_e(T)]^{-1}$; at slab center $P_e(T) \approx [2P_e(\frac{1}{2}T)]_\infty$, where the subscript denotes the escape probability from the indicated depth in a semi-infinite slab, and the factor of two accounts for losses through both faces. To obtain an order-of-magnitude estimate from equation (11-39), we replace $S_l(\tau)$ with $S_{\max} = S(\frac{1}{2}T)$ and assume εB is constant so that $\langle N \rangle \approx S_{\max}/(\varepsilon B)$. Then, using (11-24) and (11-25) to calculate $P_e(T)$, and again ignoring numerical factors of order unity, we find

$$S_{\max} \approx \varepsilon T(\ln T)^{\frac{1}{2}} B \qquad \text{(Doppler)} \qquad (11\text{-}40a)$$

$$S_{\max} \approx \varepsilon (T/a)^{\frac{1}{2}} B \qquad \text{(Voigt)} \qquad (11\text{-}40b)$$

and

$$S_{\max} \approx \varepsilon T^{\frac{1}{2}} B \qquad \text{(Lorentz)} \qquad (11\text{-}40c)$$

The behavior described above is seen very clearly in the numerical results (**53**) shown in Figure 11-3, which gives $S_l(\tau)$ for lines with Doppler profiles ($a = 0$), for various values of ε, in an atmosphere with $T = 10^4$. The dashed curve gives the solution for a semi-infinite atmosphere with $\varepsilon = 10^{-4}$. It can be seen that for $\varepsilon \geqslant 10^{-4}$ the solutions closely resemble semi-infinite

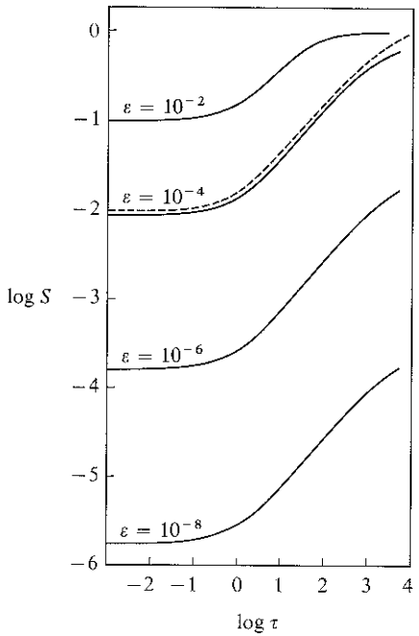

FIGURE 11-3
Line source functions in finite atmospheres with (total thickness) $T = 10^4$ and $B = 1$, for a line with a pure Doppler profile ($a = 0$). Dashed curve corresponds to the semi-infinite atmosphere solution with $\varepsilon = 10^{-4}$. Note that $S_l(\tau)$ is symmetric about $T/2$ (not shown owing to use of a logarithmic scale for abscissa. From (**53**), by permission.

atmosphere solutions at the corresponding ε, while for $\varepsilon < 10^{-4}$ the atmosphere becomes effectively thin and S_l falls below the corresponding semi-infinite-case curve, and, in fact, scales linearly with ε. Emergent intensities (**53**) are shown in Figure 11-4 for lines with $\varepsilon = 10^{-4}$, and for various values of a and T. For $T = \infty$, an absorption line is obtained in every case, with central intensity independent of a. In finite atmospheres, emission lines are obtained, for the line wing becomes completely transparent for sufficiently large x, and the intensity must go to zero. At smaller x the intensity rises rapidly, and, for effectively thick atmospheres, saturates to the semi-infinite atmosphere value. Finally, in the line core, scattering leads to a self-reversal. The profiles shown in Figure 11-4 strongly resemble those from laboratory emission sources with saturated lines, and from hot chromospheric layers above a relatively cool photosphere.

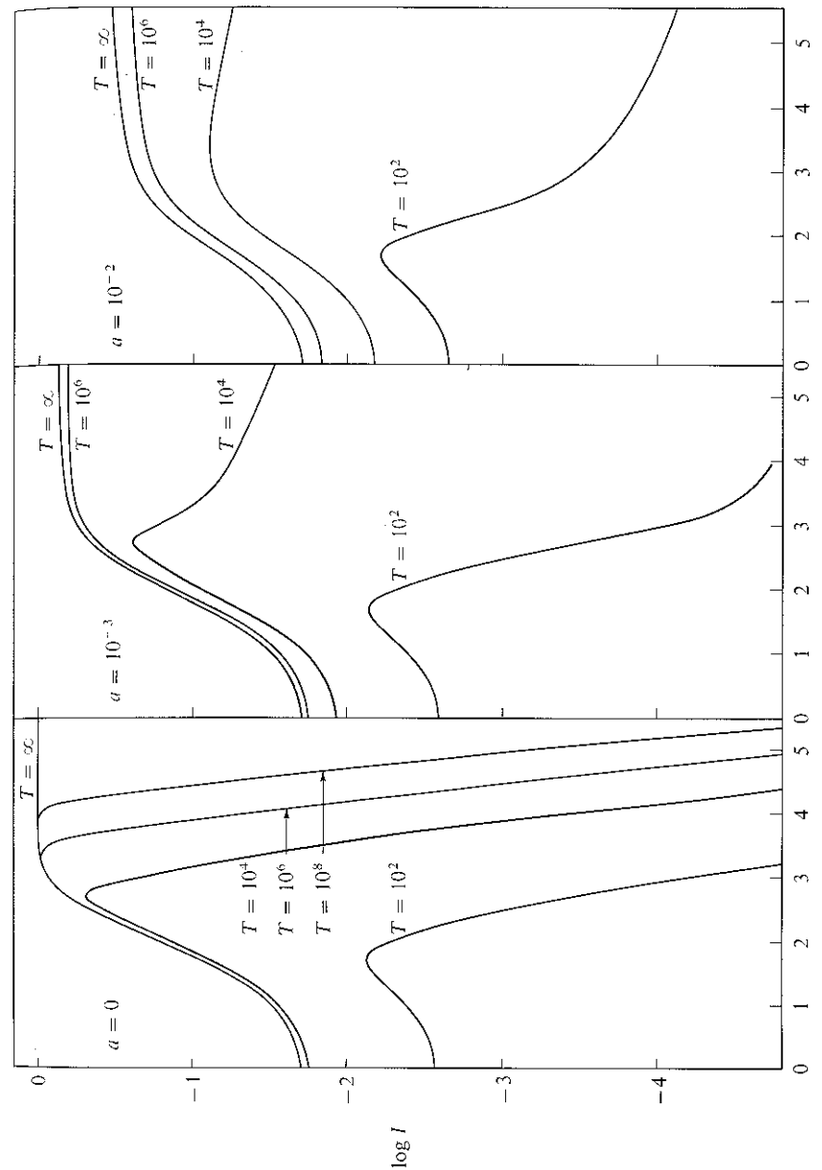

FIGURE 11-4
Emergent intensity from finite and semi-infinite atmospheres in Doppler and Voigt profiles with $\varepsilon = 10^{-4}$. From (**53**), by permission.

THE EFFECTS OF AN OVERLAPPING CONTINUUM

All of the discussion presented thus far in this section has been based on the assumption that the only opacity source is the line itself. In virtually all situations of astrophysical interest, however, the line will be superposed on a background continuum arising from bound–free absorption by other levels of the atom, or by other atomic species. As was noted in §11-1, the presence of a background continuum affects the destruction probability P_d, and sets an upper bound on the photon mean-free-path and destruction length. We therefore expect the thermalization depth to have an upper bound set by continuum processes. Furthermore, the continuum provides an additional source of emission into the line. A detailed study of the line-formation problem with a background continuum has been presented in (**315**) for idealized problems similar to those considered above in the no-continuum case; we shall summarize some of the results of this work in the following paragraphs.

For simplicity we shall consider a semi-infinite atmosphere, and assume depth-independence for the line and continuum opacities χ_l and χ_c, the line profile ϕ_x, the line thermalization parameter ε, and the Planck-function B (all of these assumptions may be dropped if numerical techniques such as the Feautrier or Rybicki methods are employed). The transfer equation is

$$\mu(dI_x/dz) = -(\chi_c + \chi_l\phi_x)I_x + \chi_l\phi_x S_l + \chi_c B \tag{11-41}$$

Defining the optical depth in terms of the line opacity, $d\tau \equiv -\chi_l\, dz$, which implies that the continuum is regarded as a perturbation, and writing $r \equiv \chi_c/\chi_l$, equation(11-41) becomes

$$\mu(dI_x/d\tau) = (\phi_x + r)(I_x - S_x) \tag{11-42}$$

where
$$S_x \equiv \frac{(1-\varepsilon)\phi_x}{(\phi_x + r)}\int_{-\infty}^{\infty}\phi_{x'}J_{x'}\,dx' + \frac{(\varepsilon\phi_x + r)}{(\phi_x + r)}B$$

$$= (1 - \xi_x)\int_{-\infty}^{\infty}\phi_{x'}J_{x'}\,dx' + \xi_x B \tag{11-43}$$

and
$$\xi_x \equiv (\varepsilon\phi_x + r)/(\phi_x + r) \tag{11-44}$$

It is easy to see that ξ_x is the *total destruction probability* for a photon at frequency x, for it is the sum of (a) the probability that a photon is absorbed in the line, $\phi_x/(\phi_x + r)$, multiplied by ε, the probability that a line photon is destroyed by collisions, and (b) the probability that the photon is absorbed in the continuum, $r/(\phi_x + r)$, multiplied by unit probability of

destruction. The formal solution of the transfer equation (11-42) yields

$$J_x(\tau) = \frac{1}{2}\int_0^{\infty} S_x(t)E_1\left|(t-\tau)(\phi_x + r)\right|(\phi_x + r)\,dt$$

$$= \frac{1}{2}\phi_x\int_0^{\infty} S_l(t)E_1\left|(t-\tau)(\phi_x + r)\right|dt + \frac{1}{2}rB\int_0^{\infty}E_1\left|(t-\tau)(\phi_x + r)\right|dt \tag{11-45}$$

The solution given by equation (11-45) could be substituted into (11-43) to produce an integral equation for the total source function $S_x(\tau)$, but the resulting equation would be two-dimensional (τ and x), and would be impractical to solve. Alternatively we can write an equation for $S_l(\tau)$; if the line source function is known, we can compute $S_x(\tau) = [\phi_x S_l(\tau) + rB]/(\phi_x + r)$ as desired. To obtain the equation for $S_l(\tau)$, the integral $\int \phi_x J_x\,dx$ is calculated from equation (11-45) and the result substituted into equation (11-11) to give

$$S_l(\tau) = (1-\varepsilon)(1-\delta)\int_0^{\infty} S_l(t)K_{1,r}\left|t-\tau\right|dt + (1-\varepsilon)\,\delta B\int_0^{\infty}L_{1,r}\left|t-\tau\right|dt + \varepsilon B \tag{11-46}$$

where
$$\delta \equiv r\int_{-\infty}^{\infty}\phi_x(\phi_x + r)^{-1}\,dx \tag{11-47}$$

$$K_{1,r}(s) \equiv \frac{1}{2}(1-\delta)^{-1}\int_{-\infty}^{\infty}E_1[(\phi_x + r)s]\phi_x{}^2\,dx \tag{11-48}$$

and
$$L_{1,r}(s) \equiv \frac{1}{2}r\,\delta^{-1}\int_{-\infty}^{\infty}E_1[(\phi_x + r)s]\phi_x\,dx \tag{11-49}$$

The particular choice of coefficients in front of the integrals defining the two kernel functions was made to assure *normalization* of the kernels.

Exercise 11-7: (a) Derive equation (11-46). (b) Show that $\lim_{r\to 0}(\delta/r) = \infty$, $\lim_{r\to 0}(1-\delta) = 1$, and hence $\lim_{r\to 0}L_{1,r}(\tau) = 0$ and $\lim_{r\to 0}K_{1,r}(\tau) = K_1(\tau)$ where $K_1(\tau)$ is defined by equation (11-19). (c) Show that $K_{1,r}$ and $L_{1,r}$ are normalized such that $\int_0^{\infty}K_{1,r}(\tau)\,d\tau = \frac{1}{2}$ and $\int_0^{\infty}L_{1,r}(\tau)\,d\tau = \frac{1}{2}$. (d) In view of (c), show that, for $\tau \to \infty$, where $S_l(\tau)$ is slowly varying and hence can be removed from the integral, $S_l(\tau) \to B$.

The physical significance of equation (11-46) becomes clear if we introduce the *average destruction probability*

$$\bar{\xi} \equiv \int_{-\infty}^{\infty}\phi_x\xi_x\,dx = \delta + \varepsilon(1-\delta) \tag{11-50}$$

for then $(1 - \bar{\xi}) = (1 - \varepsilon)(1 - \delta)$ and we can rewrite equation (11-46) as

$$S_l(\tau) = (1 - \bar{\xi}) \int_0^\infty S_l(t) K_{1,r}\,|t - \tau|\,dt + (\bar{\xi} - \varepsilon) B \int_0^\infty L_{1,r}\,|t - \tau|\,dt + \varepsilon B$$

$$\equiv (1 - \bar{\xi}) \int_0^\infty S_l(t) K_{1,r}\,|t - \tau|\,dt + \bar{\xi} B_{\text{eff}}(\tau) \qquad (11\text{-}51)$$

In the first form of equation (11-51), the first term obviously represents scattering in the line, with a total photon destruction probability given by $\bar{\xi}$. The second term represents photons fed into the line from the continuum, and the third term represents photons created by collisional processes. As in the pure-line case, the depth-variation of the source function will be determined by the properties of the kernel functions, and thermalization properties of S_l can be derived from a mathematical analysis of the asymptotic forms of $K_{1,r}$ and $L_{1,r}$ (**315**). On the other hand, the second form of equation (11-51) shows that the problem with a continuum is of the same form as the pure-line case, except that the coupling constant between the radiation and the thermal pool is $\bar{\xi}$, not ε. Indeed, exploiting the same line of reasoning as before, but using $\bar{\xi}$ for P_d, one finds that *the thermalization depth is again given by equations* (11-26), but with ε replaced by $\bar{\xi}$. Values of $\bar{\xi}$ and Λ, for $\varepsilon = 10^{-6}$ and for various values of r and a, are presented in Table 11-1. The striking result seen there is that $\bar{\xi}$ can *greatly exceed* ε, even *when* $r \leqslant \varepsilon$, with a consequent dramatic decrease in Λ. These results show that by inhibiting photon transfer in the line wings, the continuum can dominate line thermalization. Note also that whenever $r \geqslant \varepsilon$, then $\bar{\xi}$ is already sufficiently large to assure that $\Lambda < r^{-1}$, as it must be, because we know from physical considerations than over distances greater than r^{-1} thermalization

TABLE 11-1
Average Photon Destruction Probability $\bar{\xi}$ *and Thermalization Depth* Λ *for a Line with Overlapping Continuum* ($\varepsilon = 10^{-6}$)

	a					
	0		10^{-3}		10^{-2}	
r	$\bar{\xi}$	Λ	$\bar{\xi}$	Λ	$\bar{\xi}$	Λ
0	10^{-6}	1×10^{6}	10^{-6}	10^{9}	10^{-6}	10^{10}
10^{-7}	1.79×10^{-6}	5.6×10^{6}	1.87×10^{-3}	2.9×10^{6}	5.71×10^{-5}	3.1×10^{7}
10^{-6}	8.26×10^{-6}	1.2×10^{5}	5.72×10^{-5}	2.1×10^{5}	1.78×10^{-4}	3.2×10^{5}
10^{-5}	6.69×10^{-5}	1.5×10^{4}	1.85×10^{-4}	2.9×10^{4}	5.63×10^{-4}	3.2×10^{4}
10^{-4}	5.85×10^{-4}	1.7×10^{3}	7.84×10^{-4}	1.6×10^{3}	1.83×10^{-3}	3.0×10^{3}
10^{-3}	4.98×10^{-3}	2.0×10^{2}	5.23×10^{-3}	3.7×10^{1}	7.25×10^{-3}	1.9×10^{2}

SOURCE: (**315**), by permission.

in the continuum will occur; equations (11-26) with ε replaced by $\bar{\xi}$ are thus consistent.

To estimate the source function at the surface we can again use the similarity of equation (11-51) to (11-18) to infer that there exists a relation of the form

$$S_l(0) = \bar{\xi}^{\frac{1}{2}} \bar{B}_{\text{eff}} \qquad (11\text{-}52)$$

where $\bar{B}_{\text{eff}}$ is an appropriate average of $B_{\text{eff}}(\tau)$ over a thermalization length. It is apparent from the discussion given above that whenever $\bar{\xi}$ exceeds ε (either because $r > \varepsilon$ or because extensive line wings assure that $\bar{\xi} > \varepsilon$ even though $r < \varepsilon$), then both the destruction and creation terms (i.e., B_{eff}) are dominated by the continuum, and the line is said to be *continuum controlled.* This implies that $S_l(0)$, in particular, is set by the continuum even though the line opacity and line source function in the line core exceed the continuum terms. This is yet another manifestation of the fact, noted earlier, that the surface value of the source function is controlled by the line wings.

Numerical solutions for $S_l(\tau)$, in a semi-infinite atmosphere with $B = 1$ and $\varepsilon = 10^{-6}$, obtained using the method of discrete ordinates (**315**), are shown in Figures 11-5 and 11-6. It is easily seen that even modest values of

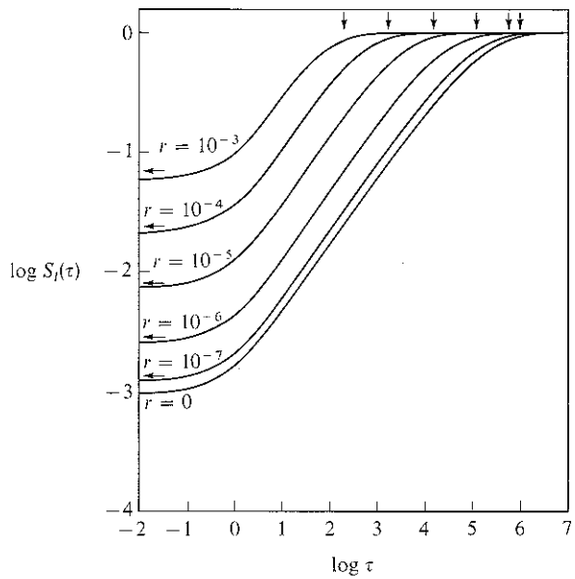

FIGURE 11-5
Line source functions in a semi-infinite atmosphere for a two-level atom with $\varepsilon = 10^{-6}$ and a pure Doppler profile ($a = 0$), with an overlapping continuum having various values of r. From (**315**), by permission.

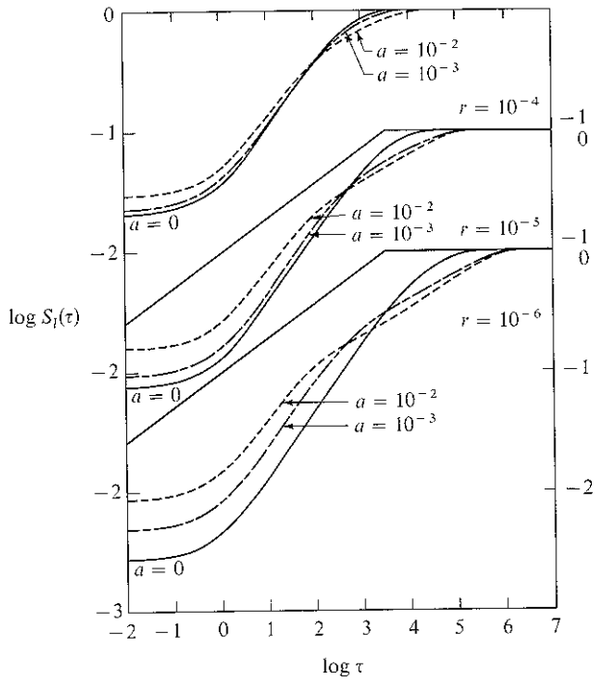

FIGURE 11-6
Line source functions in a semi-infinite atmosphere for a two-level atom with $\varepsilon = 10^{-6}$ and an overlapping continuum with various values of r, showing effects of different line profiles. From (**315**), by permission.

r produce major changes in the source function; the effects are even more striking in finite slabs (**315**). The results in Figure 11-5 pertain to a pure Doppler profile; the vertical arrows indicate the predicted thermalization depths listed in Table 11-1, and the horizontal arrows show $\overline{\zeta}^{\frac{1}{2}}$, the value of $S_l(0)$ if $B_{\text{eff}} \equiv 1$. For Voigt profiles (Figure 11-6) a fairly good estimate for $S_l(0)$ is obtained if one adopts

$$\bar{B}_{\text{eff}} = (\Lambda/\alpha)^{-1} \int_0^{(\Lambda/\alpha)} B_{\text{eff}}(\tau)\, d\tau \tag{11-53}$$

with α chosen to be a little larger than 3; this choice allows for the strong decrease in the true weighting function with depth. For *finite* slabs, a first approximation to $S_{\max}$ at slab center is given by equations (11-40) with ε replaced by $\overline{\zeta}$; this accounts for photons emitted by the continuum.

In real stellar atmospheres the effects of a background continuum on line-formation and thermalization are usually major. For early-type stars

only the strongest lines have values of r as small as 10^{-4} or 10^{-3}, and continuum terms therefore are generally dominant. On the other hand, the H and K lines of Ca II in the solar spectrum have $r \approx 10^{-9}$, and hence their thermalization is essentially independent of the overlapping continuum.

EFFECTS OF DEPTH-VARIABLE THERMALIZATION PARAMETERS AND LINE PROFILES

In the idealized problems considered thus far in this section it has been assumed, for the sake of simplicity, that all of the quantities B, r, ε, and ϕ_ν are independent of depth, whereas in real stellar atmospheres they may all be strong functions of depth. These variations can, of course, markedly affect the line-formation process, the behavior of the source function, and computed line profiles. To account for these complications in attempts to model accurately real stellar spectra, recourse must be had to direct numerical methods such as those described earlier in this chapter. But physical insight can be gained from a study of simple parameterizations of the variations of some of the quantities mentioned above. A discussion of the extremely important effects of Planck function variations will be deferred to §11-3, where line-formation in the presence of a chromospheric temperature rise will be described. Variations in the parameter r can be important, but are not easy to summarize in a few words; see (**18**, Chap. 3) for further discussion. We shall concentrate here on the effects of depth-variations of ε and ϕ_ν.

The collisional de-excitation parameter ε is proportional to the electron density, and hence must reflect the rise of the density with depth in the atmosphere. If the atmosphere is supposed to be essentially isothermal, then hydrostatic equilibrium implies that the total density increases linearly with m, the column-mass; if further we suppose that (χ/ρ) is essentially constant, then τ is also proportional to m. Then in an early-type star where the material is appreciably ionized, we expect n_e to be proportional to τ; in a later-type star the material may be neutral in the outer layers and may then abruptly ionize at some depth, in which case a much more sudden rise in n_e (and ε) may occur. Even though an actual numerical solution is required to determine $S_l(\tau)$ with precision when ε is variable, it is reasonable on physical grounds that we should still expect thermalization to occur at a depth Λ where $P_e(\Lambda) \approx P_d(\Lambda)$. To illustrate the usefulness of this idea, suppose a rapid rise in ε occurs, from some value ε_1 to higher value ε_2, at a particular depth τ_0 in an atmosphere with r, B, and ϕ all constant. Then if $\tau_0 > 1/\varepsilon_1$ (for a Doppler profile) it is clear that the line will already have thermalized, and hence the rise in ε has no effect. In this case $S_l(0)$ will equal $\varepsilon_1^{\frac{1}{2}}B$. Conversely, suppose that $\tau_0 < 1/\varepsilon_2$. Then the line source function thermalization will have been delayed by the low value of ε at the surface to a predicted depth of $1/\varepsilon_1$; but at that depth ε already equals $\varepsilon_2 > \varepsilon_1$. Hence the line will

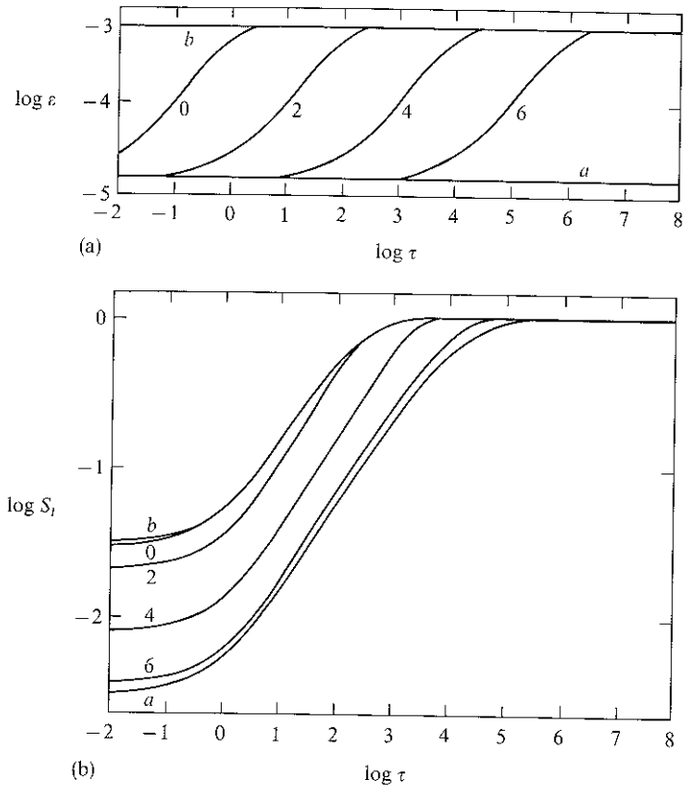

FIGURE 11-7
(a) Variation of ε with depth: curves are labeled with $\log \tau_0$, and $\varepsilon(\tau) = 10^{-3}[1 - 0.99 \exp(-\tau/\tau_0)]$. (b) Line source function in a semi-infinite atmosphere with $B = 1$, for a line with a pure Doppler profile and $\varepsilon(\tau)$ as shown in (a). Curves labeled "a" and "b" correspond to constant values of ε of 10^{-5} and 10^{-3} respectively. From (**284**, 101).

behave as if ε were constant at $\varepsilon \equiv \varepsilon_2$, and $S_l(0)$ will equal $\varepsilon_2 \frac{1}{2} B$. These qualitative expectations are confirmed by detailed calculations (**284**, 101) such as those shown in Figure 11-7 where ε is assumed to have the form $\varepsilon(\tau) = 10^{-3}[1 - 0.99 \exp(-\tau/\tau_0)]$. Notice that, when $\tau_0 > 10^5$, S_l behaves as if $\varepsilon \equiv 10^{-5}$, whereas if $\tau_0 \lesssim 10^2$, S_l behaves as if $\varepsilon \equiv 10^{-3}$.

A change in the form of the line profile alters the way photons get redistributed into the line wings, and their probability of escape. Numerical studies (**320**; **18**, 51) have revealed a number of interesting effects of profile depth-variations upon the line source function. Suppose, for example, the Doppler width rises sharply near the surface in an atmosphere with $B =$

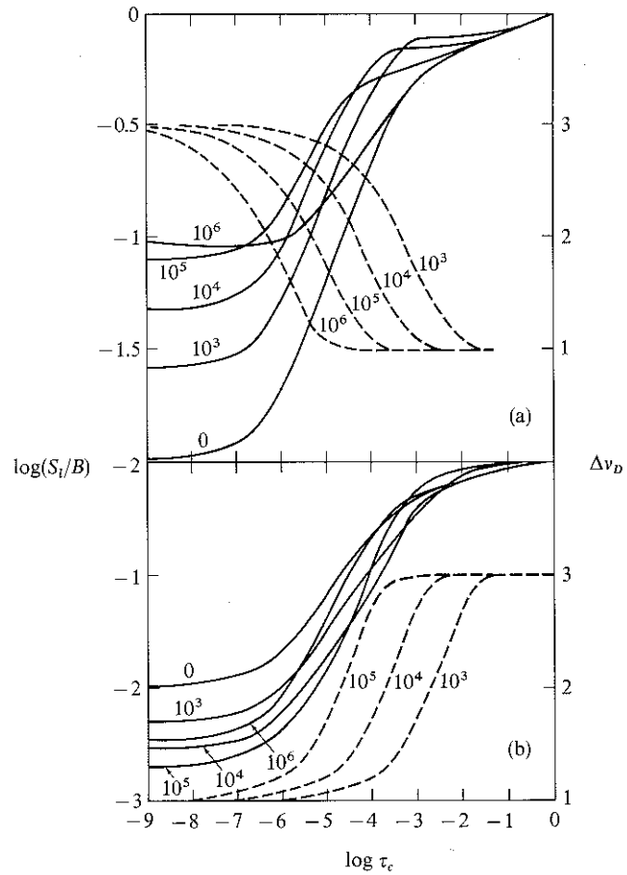

FIGURE 11-8
Line source functions in a semi-infinite atmosphere with $B = 1$, for a line with $\varepsilon = 10^{-4}$, $r = 10^{-6}$, $a = 10^{-3}$, and a depth-variable Doppler width. (a) Doppler width rising toward surface, $\Delta\nu_D(\tau) = 1 + 2 \exp(-\alpha\tau_c)$. (b) Doppler width decreasing toward surface, $\Delta\nu_D(\tau) = 3 - 2 \exp(-\alpha\tau_c)$. Curves are labeled with parameter α. Curves labeled "0" have *constant* Doppler width. *Ordinates*: Left-hand scale and *solid* curves give $\log(S_l/B)$; right-hand scale and *dashed* curves give $\Delta\nu_D$. From (**18**, 52), by permission.

constant (see Figure 11-8a). Then we find that the surface value $S_l(0)$ is increased markedly because the broader line-wings can now intercept radiation from deeper layers. As the position of the rise in $\Delta\nu_D$ moves deeper in the atmosphere, the surface effect diminishes, basically because the surface layer becomes more and more opaque in the wings, and the bright underlying

continuum radiation becomes more attenuated. At the same time, the source function at great depth also increases, basically because the higher opacity in the line-wings in the surface layer impedes the escape of radiation from below; this is sometimes referred to as the *reflector effect*. The effect at depth is greater, the thicker the upper layer (until, of course, the rise lies below the point of thermalization). Notice in Figure 11-8a that when $\alpha = 10^6$ the upper layer is optically thin, and thus there is essentially no reflector effect even though the surface rise is largest.

When the Doppler width decreases sharply at the surface, the value of $S_l(0)$ drops because the narrower profile no longer intercepts bright radiation in its wings. Radiation trapping is also reduced in lower layers, and hence S_l decreases there as well. At very great depth, however, the effect is the opposite, and S_l actually rises a little above its value for constant $\Delta\nu_D$, because now the decreased bandwidth available to the lines implies a decreased escape probability.

It must also be noted that, in addition to changing $S_l(\tau)$, variations in the absorption profile change the run of optical depth, at a given frequency, with physical depth. In view of the Eddington–Barbier relation, it is obvious that this implies a change in the "mapping" of S_l into $I_\nu(0)$, and hence in the emergent intensity distribution in the line profile, and its center-to-limb behavior [see, e.g., (**26**)].

11-3 The Two-Level Atom with Continuum

THE SOURCE FUNCTION

The discussion given in the preceding sections of this chapter is based on a very schematic and admittedly restrictive atomic model. Naturally the true situation for any real atom is more complicated. To introduce some of the physically important effects while retaining analytical simplicity, let us now consider an atomic model consisting of two bound states and a continuum. This model provides at least a rough representation of reality for resonance lines and also for subordinate lines when the resonance lines are in radiative detailed balance. The addition of the continuum greatly increases the number of processes that may take place. In the strict two-level case, the only processes that can occur are photoexcitations or collisional excitations from the lower to upper state, and their inverses. Now there are, in addition, photoionizations and collisional ionizations from the bound levels to the continuum, and radiative and collisional three-body recombinations to each bound level. It is clear that this is a much more general model, and we shall find that the additional physics has major implications for the line-formation process.

As before, assume complete redistribution, so that S_l is given by equation (11-4). The statistical equilibrium equations now are

$$n_l\left(B_{lu}\int\phi_\nu J_\nu\,d\nu + C_{lu} + R_{l\kappa} + C_{l\kappa}\right) = n_u\left(A_{ul} + B_{ul}\int\phi_\nu J_\nu\,d\nu + C_{ul}\right) + n_l^*(R_{\kappa l} + C_{l\kappa}) \quad (11\text{-}54)$$

and

$$n_u\left(A_{ul} + B_{ul}\int\phi_\nu J_\nu\,d\nu + C_{ul} + R_{u\kappa} + C_{u\kappa}\right) = n_l\left(B_{lu}\int\phi_\nu J_\nu\,d\nu + C_{lu}\right) + n_u^*(R_{\kappa u} + C_{u\kappa}) \quad (11\text{-}55)$$

for the lower and upper levels, respectively. The photoionization and recombination rates are given by equations (5-66) and (5-67). Solving equations (11-54) and (11-55) for (n_l/n_u), substituting into equation (11-4), and making use of the Einstein relations we obtain

$$S_l = \left[\int\phi_\nu J_\nu\,d\nu + \varepsilon B_\nu(T) + \theta\right]\Big/(1 + \varepsilon + \eta) \quad (11\text{-}56)$$

where ε is given by equation (11-7),

$$\eta \equiv \frac{(R_{u\kappa} + C_{u\kappa})n_l^*(R_{\kappa l} + C_{l\kappa}) - g_l(R_{l\kappa} + C_{l\kappa})n_u^*(R_{\kappa u} + C_{u\kappa})/g_u}{A_{ul}[n_l^*(R_{\kappa l} + C_{l\kappa}) + n_u^*(R_{\kappa u} + C_{u\kappa})]} \quad (11\text{-}57)$$

and

$$\theta \equiv \left(\frac{2h\nu^3}{c^2}\right)\left(\frac{g_l}{g_u A_{ul}}\right)\frac{(R_{l\kappa} + C_{l\kappa})n_u^*(R_{\kappa u} + C_{u\kappa})}{[n_l^*(R_{\kappa l} + C_{l\kappa}) + n_u^*(R_{\kappa u} + C_{u\kappa})]} \quad (11\text{-}58)$$

Equation (11-56) was first derived by Thomas (**622**) as an extension of earlier work by Milne (**416**, 159–164) and Strömgren (**613**), and was studied extensively in an important series of papers by Jefferies and Thomas (**335**; **336**; **623**; **337**).

Exercise 11-8: Derive equations (11-56) through (11-58) in detail.

Despite the apparent complexity of these expressions, each term admits a simple interpretation. Consider the numerator of equation (11-56). The first term again represents the scattering reservoir. The second term is the thermal source, giving the rate at which photons are created by collisional excitation; note that this term depends upon the local value of T_e, the *electron kinetic temperature*. The third term is proportional to the total rate at which electrons are ionized from the ground state to the continuum,

times the fraction that recombines to the upper state and thus becomes available for emission into the line by radiative decay to the ground state. Similarly, in the denominator, the second term accounts for photons destroyed by collisional de-excitation of the upper state. The third term is a new sink term that is proportional to the total ionization rate from the upper state times the fractional recombination rate to the lower state; this term clearly accounts for the destruction of potential line photons via the continuum processes. We thus find all of the basic physics of the situation represented clearly in the source function.

Further insight follows from rewriting θ as ηB^*, where B^* is found to be

$$B^* = \left(\frac{2h\nu^3}{c^2}\right)\left\{\left(\frac{n_l^* g_u}{n_u^* g_l}\right)\left[\frac{(R_{u\kappa} + C_{u\kappa})(R_{\kappa l} + C_{l\kappa})}{(R_{l\kappa} + C_{l\kappa})(R_{\kappa u} + C_{u\kappa})}\right] - 1\right\}^{-1} \quad (11\text{-}59)$$

Exercise 11-9: Verify equation (11-59).

It is clear that B^* bears a formal resemblance to the Planck function, and, in fact, we can think of it as $B^* = B_\nu(T_r)$, where T_r is a characteristic *radiation temperature* whose value is set by the photoionization and recombination rates in the two bound–free continua. This radiation temperature can, in general, be quite different from T_e. At great depth in the atmosphere where the continua are optically thick, $J_\nu \to B_\nu(T_e)$; then $R_{l\kappa} \to R^*_{l\kappa} = R^*_{\kappa l}$, and $R_{u\kappa} \to R^*_{u\kappa} = R^*_{\kappa u}$, so that $B^* \to B_\nu(T_e)$, as expected, and $S_l \to B_\nu(T_e)$. But near the surface, the continua may become transparent (while the line remains optically thick) and the continuum radiative rates become essentially fixed. Depending on the relation of the radiation temperature (which characterizes J_ν emerging from $\tau_c \sim 1$) to the local electron temperature, quite distinct situations emerge. For example, suppose that collisions are negligible, and that $R_{l\kappa} > R_{\kappa l}$ while $R_{\kappa u} > R_{u\kappa}$ (i.e., suppose that a relatively cool medium is irradiated by a "hot" radiation field in the ground-state continuum). Then it is clear that $B^* > B_\nu(T_e)$, and that S_l will be larger than the value it would have had by coupling to the thermal pool; the excess emission comes from a preferential depopulation of the lower state into the upper via the continuum, followed by radiative de-excitation in the line. If the line is an absorption line, an increase in S_l tends to *weaken* it; taken to the extreme limit, this is the mechanism that produces the photoionization–recombination–cascade emission spectrum in a nebula. If the inequalities posed above are reversed, then the upper state is selectively depopulated into the lower, and S_l decreases; if the line is an absorption line, it will *strengthen*.

The two essential points that have emerged from the above analysis are the following. (a) The source function for a given line contains terms in *other* transitions (the two continua for the present model). This result is quite generally true and carries over to the multilevel case (see §12-1). (b) The new terms do not couple S_l to the local thermal pool, characterized by T_e, as do collisions, but rather to a radiation temperature that may be markedly hotter or cooler than the local kinetic temperature.

CLASSIFICATION OF LINES

In view of the fundamentally differing natures of the various source and sink terms that appear, we may expect that line-transfer problems may show substantially different characteristics depending on which terms provide the dominant contributions to S_l. Thomas has suggested (**622**; **626**, 174) a classification of lines into broad categories by consideration of which terms are the largest for different atoms in typical stellar-atmosphere regimes. For example, if $\varepsilon > \eta$ and $\varepsilon B_\nu > \eta B^*$, the line is called *collision dominated*; here S_l couples to T_e. On the other hand, if $\eta > \varepsilon$ and $\eta B^* > \varepsilon B$, the line is *photoionization dominated*, and the line couples to a characteristic $T_r \neq T_e$. Intermediate cases where, say, $\varepsilon > \eta$ but $\eta B^* > \varepsilon B$, or vice versa, are called *mixed domination* lines. The recognition of these classes of lines represents a considerable advance over the rather ill-defined classical division of lines into "absorption" and "scattering" categories, and has led to important insights about line-formation. In particular, the emergent profiles for lines in these classes are quite different when a chromospheric temperature rise occurs in the outer layers of the atmosphere.

The category to which any particular line belongs depends (via the atomic cross-sections involved) upon the structure of the ion from which it arises, and upon the structure of the atmosphere (because of the dependence of the relevant rates upon atmospheric parameters such as temperature, density, and incident radiation fields). Different lines of the same ion will, in general, fall into different classes, and a comprehensive a priori classification is not possible; rather, each case must be examined in turn. Broad groupings of lines in a solar-type atmosphere have been suggested by Thomas; these are displayed in Table 11-2, and can be understood qualitatively as follows. The resonance lines of the singly ionized metals are collision dominated because the excitation energy is only a few electron volts (compared to a thermal

TABLE 11-2
Categories of Line Source Functions in a Solar-Type Atmosphere

Collision dominated	Photoionization dominated
Resonance lines of singly ionized metals (Mg^+, Ca^+, Sr^+, etc.)	Resonance lines of neutral metals
Resonance lines of H and other nonmetals (C, N, O, etc.)	Hydrogen Balmer lines

SOURCE: Adapted from (**623**; **626**, 174).

energy of about 0.5–1 eV), while the ionization energies are 11–15 eV; recalling that both the collision and photoionization rates scale as $\exp(-E_0/kT)$, it is plausible that the collisions should prevail. Even for hydrogen, where E_0 in $L\alpha$ is 10 eV, the collision rates dominate because the solar u.v. radiation is weak and the photoionization rate is small. In contrast, for the hydrogen Balmer lines and the neutral metals, the relevant photoionization edges lie in regions where the solar radiation field is intense, and the photoionization rates prevail over collision rates. The sodium D-lines are an exception because of an unusually large collision cross-section (**284**, 333; **284**, 347); this example serves as a warning that the broad classes listed are *illustrative*, and detailed analysis is needed in each case. In a higher temperature regime (e.g., in an O-star) some of the remarks just made are no longer valid, and lines switch categories. For example, the radiation field in the Lyman continuum becomes extremely intense (the hydrogen is virtually completely ionized), and the hydrogen Lyman lines become photoionization dominated. The higher subordinate lines now have energies $E_0 \lesssim kT$ ($\sim$3 eV at O-star temperatures), and are relatively weak, hence are formed in deeper, denser layers; this tends to lead to collision domination of the subordinate series.

LINE-FORMATION IN THE PRESENCE OF A CHROMOSPHERE

The great physical importance of the division of lines into the two broad categories described above becomes strikingly apparent when we consider the nature of the source function variation, and emergent line profile, in an atmosphere with a chromospheric temperature rise outward. The basic point is that collision-dominated lines are linked to the *local electron kinetic temperature*, while photoionization-dominated lines are not. The latter couple instead to a *radiation temperature* characteristic of the energy distribution (emitted at some other point in the atmosphere) in the continua of the upper and lower levels. We therefore expect collision-dominated lines to be at least partly responsive to the local temperature, and to exhibit, in their profiles, features attributable to the outward temperature rise. On the other hand, we expect photoionization-dominated line-profiles to be insensitive to variations in local parameters (in particular, to variations in T_e). The strong dichotomy of behavior just described was clearly demonstrated in the pioneering work of Jefferies and Thomas (**336**). They showed that these considerations explained the observed presence of emission cores in lines such as the H and K lines of Ca^+ (and the analogous lines of Mg^+), and their simultaneous absence in the hydrogen Balmer lines.

In their work Jefferies and Thomas adopted a schematic continuum source function of the form

$$S_c(\tau) = B_\nu(T_e) = S_1(1 + \alpha\tau_c + \beta e^{-\gamma\tau_c}) \qquad (11\text{-}60)$$

which provides an approximate representation of a temperature distribution that has: a uniform gradient in the photosphere; a plateau at a minimum temperature at some characteristic depth (they used $T_{\min} \approx 4000°$K at $\tau_c \approx 10^{-2}$); and then a steep outward gradient mimicking the sharp chromospheric rise. The transfer equation was solved in the Eddington approximation:

$$\frac{1}{3}(d^2J_\nu/d\tau^2) = (\phi_\nu + r)^2[J_\nu - (\phi_\nu + r)^{-1}(rS_c + \phi_\nu S_l)] \qquad (11\text{-}61)$$

using the method of discrete ordinates, for a Doppler profile and for typical values of r, ε, η, and B^*.

Results for a collision-dominated line ($\eta \equiv 0$) with $\varepsilon = 10^{-4}$, $r = 10^{-4}$, and various choices for the continuum source function are shown in Figure 11-9. Note that at great depth S_l thermalizes to S_c, but as the wings begin to become transparent ($\tau_c \lesssim 1$), S_l drops below S_c. Proceeding outward, S_c rises very steeply at $\tau_c \approx 10^{-2}$. The line source function tries to follow this rise through collisional coupling, and actually does increase outward, but ultimately the effects of scattering dominate, and, at the surface, S_l lies about three orders of magnitude below $B_\nu(T_e)$. It follows from the Eddington–Barbier relation that the depth-variation of $S_l(\tau)$ should be reflected in the frequency-variation of the emergent intensity. The computed emergent

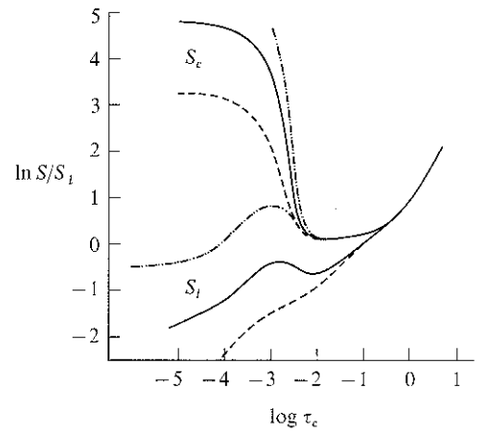

FIGURE 11-9
Line source functions for a collision-dominated line in a semi-infinite atmosphere with a chromospheric temperature rise. Upper curves show continuum source function S_c, and lower curves show corresponding line source functions S_l. In all cases $\varepsilon = 10^{-4}$ and $r = 10^{-4}$. From (**336**) by permission.

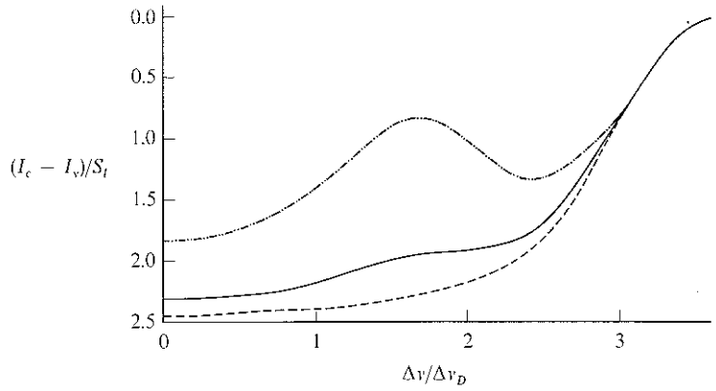

FIGURE 11-10
Line profiles for a collision-dominated line in a semi-infinite atmosphere with a chromospheric temperature rise. Curves are coded to correspond to those in Figure 11-9. From (**336**), by permission.

intensity is shown in Figure 11-10; only half of the profile (which is symmetric about line-center) is shown. Note that the lines show an emission peak near ± 1.5 Doppler widths, with a central absorption reversal; lines of this form are called *doubly reversed*. The overall variation of intensity within the line is in good *qualitative* agreement with the observed behavior in the solar Ca II H- and K-lines and the corresponding lines of Mg II. Numerous other calculations, with a chromospheric temperature rise, for other choices (sometimes including depth-variations) of ε, a, r, ϕ_ν, and different parameterizations of $B_\nu(\tau)$ can be found in the literature [see, e.g., (**284**, 101; **26**; **27**; **18**, 45–48)]. All show the same kind of behavior as that shown in Figure 11-10, and some yield semiquantitative agreement with the solar data; excellent quantitative fits to the observations are obtained when multilevel calculations, using realistic atmospheric models and atomic models for Ca^+, are employed (see §12-1).

The behavior of photoionization-dominated lines, for the same assumed Planck function, contrasts strongly with that of collision dominated lines. Results for a line with $\varepsilon = 0$, $\eta = 10^{-2}$, and $r = 10^{-4}$ are shown in Figure 11-11. At very great depth, $S_l \to S_c$ when the line thermalizes in the continuum. Proceeding outward, S_l remains fixed at the value set by B^*, and at a line optical depth of order $1/\eta$, shows a strong drop outward, characteristic of scattering, to a value equal to $\eta^{\frac{1}{2}}B^*$. *The source function shows no response whatever to the variation of* $B_\nu(T_e)$. Note that proceeding outward, S_l first lies above S_c (because B^* has a radiation temperature set in the continua in deeper, hotter layers), and then, near the surface, lies below S_c (which rises

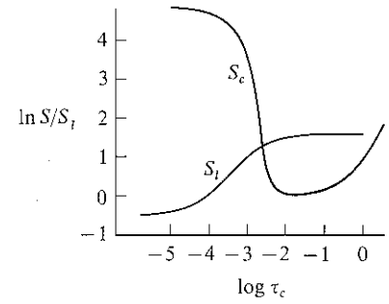

FIGURE 11-11
Source function for a photoionization-dominated line in a semi-infinite atmosphere with a chromospheric temperature rise. From (**336**), by permission.

rapidly to its chromospheric value). The line profile is a pure absorption feature with no hint of central emission reversal. This behavior is precisely what is observed for the Balmer lines and, in fact, the run of S_l correlates well with the empirical source function deduced by Athay and Thomas (**29**).

The recognition that the Ca II H- and K-lines have source functions that couple partially to the local Planck function variation led to the realization that the emission features often observed in the cores of these lines in stellar spectra contain valuable information about the temperature distribution in stellar chromospheres. If suitably analyzed, these lines provide unique diagnostic tools for the determination of the physical structure of the outer layers of solar-type stars. It is known that the intensity of the emission correlates with stellar age (**675**; **677**), and this fact offers the possibility that by combining accurate chromospheric diagnostics with a reliable theory of stellar evolution, we may be able to infer the time-evolution of the solar chromosphere. One of the most fascinating empirical results relating to chromospheric Ca II H and K-line emission is known as the *Wilson–Bappu effect* (**676**), which shows that the half-intensity-width of the emission components correlates closely with stellar luminosity, over a range of 10^6 in L. Several theoretical suggestions have been made to explain this observation, but, as was early emphasized by Jefferies and Thomas (**337**), an understanding of the phenomenon must be based on an accurate physical picture of line-formation. Their work, and more recent efforts, have elucidated the dependence of the emission intensity and width upon the form of $T_e(\tau)$ (amplitude and depth of rise) and upon the functions ε, r, and ϕ_ν.

Some useful insight concerning the basic properties of chromospheric emission lines can be gained from approximate scaling rules [(**24**; **18**, 46)] such as those summarized in Table 11-3. These give estimates of the surface value, $S_l(0)$, and peak value, $S_l(\max)$, of the source function in a chromospheric slab, with a Planck-function variation given by equation (11-60) with S_1 set to unity. The rules are based on the assumption that the Doppler width is independent of depth; Doppler width variations alter the results

TABLE 11-3
Scaling Rules for Line Source Function in Finite Chromospheric Slabs

Chromospheric thickness	Doppler profile ($a < \varepsilon$) $S_l(0)$	Doppler profile ($a < \varepsilon$) $S_l(\max)$	Voigt profile $\varepsilon < a < 1$ $S_l(0)$	Voigt profile $\varepsilon < a < 1$ $S_l(\max)$
Optically thin $(\gamma r)^{-1} < 1$	$\bar{\xi}\beta$	$\bar{\xi}\beta$	$\bar{\xi}\beta$	$\bar{\xi}\beta$
Effectively thin $1 < (\gamma r)^{-1} < \Lambda$	$\bar{\xi}\beta(\gamma r)^{-\frac{1}{2}}$	$\bar{\xi}\beta(\gamma r)^{-1}$	$\bar{\xi}\beta(\gamma r a)^{-\frac{1}{2}}$	$\bar{\xi}\beta(\gamma r a)^{-\frac{1}{2}}$
Effectively thick $\Lambda < (\gamma r)^{-1}$	$\bar{\xi}^{\frac{1}{2}}\beta$	β	$\bar{\xi}^{\frac{1}{2}}\beta$	β

SOURCE: Adapted from (**18**, 46), by permission.

somewhat (**24**) and, more importantly, strongly affect the emergent line-profile for a *given* $S_l(\tau)$. The parameters of interest for describing the run of $B_\nu(T_e)$ are β, the amplitude of the rise (assumed to be $\gg 1$), and γ^{-1}, the continuum depth at which the rise occurs [the corresponding line depth is $(\gamma r)^{-1}$]. If $(\gamma r)^{-1} \ll \Lambda$ (the thermalization depth), then S_l will respond only weakly to the chromospheric temperature rise; but if the inequality is reversed, a strong effect will occur. If the slab is *effectively thick* [i.e., $(\gamma r)^{-1} > \Lambda$, where $\Lambda = \bar{\xi}^{-1}$ or $a\bar{\xi}^{-2}$ for Doppler and Voigt profiles, respectively, and $\bar{\xi}$ is the total destruction probability given by equation (11-50)], then S_l saturates to β at depth and falls by a factor of $\bar{\xi}^{\frac{1}{2}}$ at the surface. If the slab is *optically thin*, then S_l just equals the local creation term $\bar{\xi}\beta$. If the slab is *effectively thin*, the maximum value of S_l is given by equations (11-40) with $T = (\gamma r)^{-1}$. Note that the quoted results for the Voigt profile presume that $a(\gamma r)^{-1} > 1$; if a is so small that this is not true, the line wings are negligible and the results listed for a Doppler profile apply instead. The surface value of S_l can be derived by recognizing that if $S_{\max} \propto \bar{\xi}\beta\langle N\rangle$ (where $\langle N\rangle$ gives the mean number of scatterings needed to escape), then from random-walk arguments we expect $S_l(0) \propto \bar{\xi}\beta\langle N\rangle^{\frac{1}{2}}$. Note that the effectively-thin results merge smoothly with the effectively-thick results when $(\gamma r)^{-1} = \Lambda$, and with the optically-thin results when $(\gamma r)^{-1} = 1$.

From the results of Table 11-3 we see that, with a given chromospheric structure, some collision dominated lines will respond strongly to the temperature rise, while others may not, depending on the values of ε, r, and a appropriate to them. For example, for the Ca II lines in the solar chromosphere we have (to order of magnitude only!): $a \sim 10^{-3}$, $r \sim 10^{-9}$, $\varepsilon \sim 3 \times 10^{-4}$, and $\gamma \sim 10^6$; therefore $\Lambda \sim 10^4$, while $(\gamma r)^{-1} \sim 10^3$, so the lines are effectively thin. Furthermore $a(\gamma r)^{-1} \sim 1$, so $(S_{\max}/\beta) \sim \varepsilon(\gamma r a)^{-\frac{1}{2}} \sim$ 0.3, and thus the coupling of the Ca II lines to the temperature rise is relatively inefficient. In contrast, for the Mg II resonance lines, r is about a factor of 10 smaller (because Mg is about 10 times more abundant than Ca), and the other parameters are about the same; in this case we obtain a much stronger coupling to the rise and much brighter emission, as observed. It should be noted that the arguments given about are meant only to be *qualitative*, for the depth-variation of the parameters is actually fairly complicated, and not well-represented everywhere by the values chosen; in particular, equation (11-60) provides a relatively poor fit to the actual run of $B_\nu(\tau)$. Detailed analyses of Ca II line-formation in stellar chromospheres have been performed using realistic atomic and atmospheric models (**56**; **58**), and a physically plausible explanation of the Wilson–Bappu effect has begun to emerge (**57**).

11-4 Static Extended Atmospheres

As was described in §7-6, many stars have extended atmospheres whose thicknesses are comparable to the radius of the star. We shall suppose that, to a first approximation, these atmospheres are spherically symmetric. For the purposes of the present discussion we shall consider the stellar radius r_* to be that of the surface on which $\tau_R \approx \frac{2}{3}$. We imagine that this "photospheric core" is surrounded by an envelope of large size, within which unit optical depth in the most opaque spectral regions (i.e., line centers) is encountered at radii $R \gg r_*$. In reality, virtually all stars with very extended atmospheres also have large-scale velocity fields (usually overall atmospheric expansion) that strongly affect—indeed dominate—transfer in the lines, so that the assumption of a *static* atmosphere is physically less useful in an analysis of line-formation than it was for the continuum. Nevertheless, there are some extremely important effects of a fundamentally geometric origin that enter; it is worthwhile to examine these here, and to defer a discussion of velocity-field effects until Chapter 14.

As seen by an outside observer, the *size* of the emitting surface where $\tau_\nu = 1$ at the more opaque frequencies, in particular the cores of spectral lines, can be much larger than that of the continuum. Then the line has a larger *effective emitting area*, and if we assume LTE and suppose that the envelope is essentially isothermal, it is clear that the line will appear in *emission* relative to the continuum. This behavior contrasts with the result for an isothermal planar atmosphere where the line is neither in absorption or emission. In fact, the basic geometric effect just described is actually the primary mechanism that produces the extremely intense emission in very opaque spectral lines observed—e.g., in Wolf–Rayet spectra. The assumption

that the atmosphere is essentially isothermal implies a nonradiative source of energy input, for we have seen in §7-6 that in radiative equilibrium (in LTE), $J \sim B \sim r^{-2}$ in an extended atmosphere; thus the situation just described may be a bit extreme (though we have no strong physical reason to prefer radiative equilibrium—recall the solar corona!). Also, if the line source function has a scattering term, we expect $S_l(\tau)$ to decrease outward (in fact the drop is enhanced by extension effects; see below). However, it is clear that, for any *given* $S_l(\tau)$, the change in effective emitting area from line core to wing will always tend to increase the core emission, relative to the continuum, compared to the value it would have had in a planar atmosphere. Absorption lines will tend to weaken, and emission lines will tend to grow brighter, as atmospheric size increases.

A second effect of atmospheric extension is a systematic *increase of the escape probability* at a given (radial) optical depth. It is obvious from elementary geometric considerations that along all rays, except the single ray at $\mu = +1$, the optical path-length from the test point to the surface is *smaller* in a spherically symmetric atmosphere than in a plane-parallel semi-infinite atmosphere of identical properties (i.e., same run of physical variables along the radial direction in the two cases). Thus, in view of equation (11-22), $P_e(\tau)$ increases, and we expect $S_l(\tau)$ to decrease accordingly.

A third effect, which also increases the effective probability of escape, is the systematic *bias toward larger radii* in the scattering. Suppose that the scattering process is isotropic and that the material is homogeneous, so that the photon mean-free-path is the same in all directions, and that $l = 1$ defines a spherical volume around the test point. Then, from the basic geometry of the situation, it is easy to see that, in a spherical atmosphere, *more* photons end their flights at *larger* radii (hence closer to the surface); in contrast, in a planar atmosphere the probabilities that the photon ends its flight at a greater or smaller depth (by a given amount) are *equal*. If the material has an opacity that decreases strongly outward, the bias is enormously enhanced [see Table II and Fig. 2 of (**374**)]. This bias implies that the net chance of escape is enhanced even further.

The transfer equation for the two-level-atom line-formation problem [i.e., S_l given by an equation of the form of (11-56)] in spherical geometry is easily solved by the methods described in §7-6 [see equations (7-190) through (7-208) and related discussion]. Calculations for idealized models (**374**) similar to those employed earlier in this chapter provide illustrations of the concepts outlined above. We characterize the atmosphere by its outer radius R (in units of the core radius $r_c = 1$), a total line optical depth T_l, a continuum optical depth T_c, and opacities $\chi_l \propto r^{-2}$ and $\chi_c \propto r^{-2}$. We set $\varepsilon =$ constant, and $B = 1$. Source functions for an envelope with no background continuum, surrounding an *empty* core (i.e., a nebula) are shown in Figure 11-12. The two

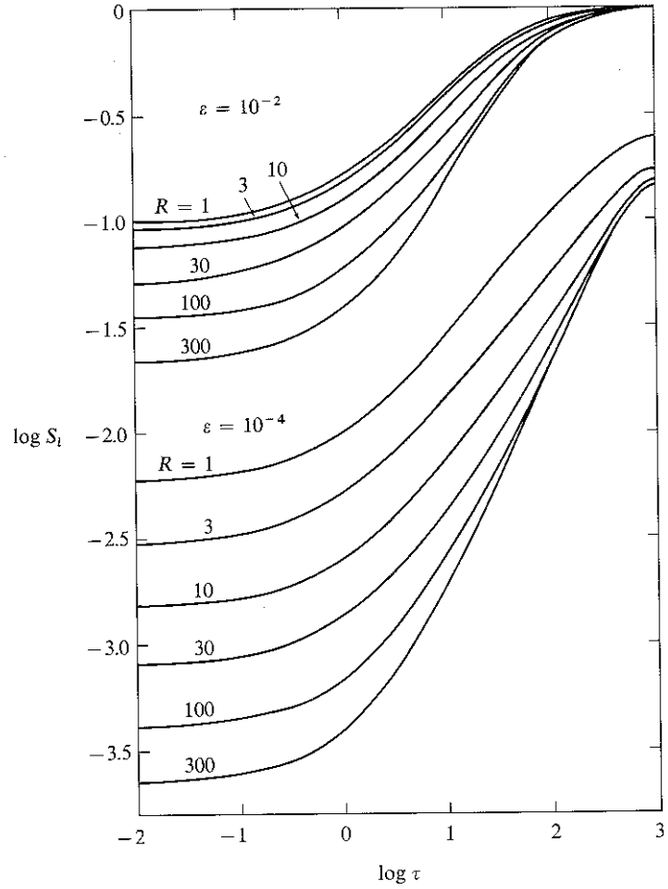

FIGURE 11-12
Line source functions in a spherical atmosphere of outer radius R (in units of $r_c = 1$), $B = 1$, $\varepsilon =$ constant, $T_c = 0$, $T_l = 10^3$, and opacity $\chi_l \propto r^{-2}$. Curves are labeled with R, and results for two values of ε, corresponding to effectively thick and effectively thin media, are shown. From (**374**), by permission.

values of ε chosen yield effectively thick and effectively thin media. The major effects of increased photon escape in decreasing S_l are readily apparent. Note that, for the effectively thick medium, the thermalization depth ($\Lambda \approx 10^2$) is scarcely affected by sphericity. Emergent flux profiles for these source functions are shown in Figure 11-13. There we see that the central reversal, which is a prominent feature in the planar limit (note *log* scale), vanishes for very

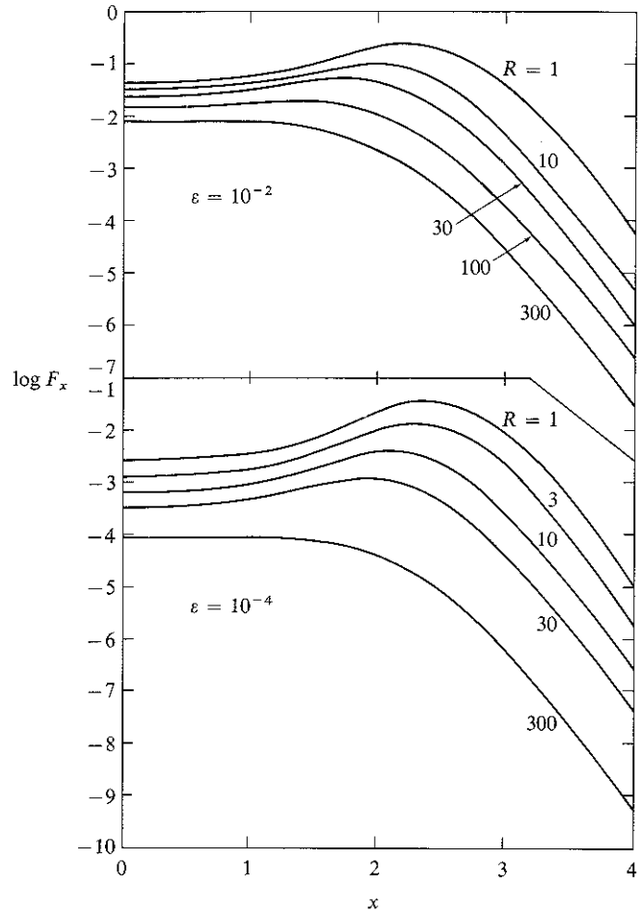

FIGURE 11-13
Emergent flux profiles from spherical atmospheres with same properties as used in Figure 11-12. From (**374**), by permission.

extended atmospheres; this is a manifestation of the greater effective emitting area in the line core. Analogous results for atmospheres with $T_c = 2$ and $\varepsilon = 2 \times 10^{-3}$, and otherwise identical to those mentioned above, are shown in Figures 14-11 (source functions; note only curves with $v_{\max} = 0$) and 11-14 (line profiles). Again, the source functions are dramatically reduced by increased escapes, and the line profile transforms from a pure absorption line to an emission line with a central reversal (the reversal results only because S_l has decreased so much at $\tau = 1$; if S_l remained unchanged, the whole line would come into emission).

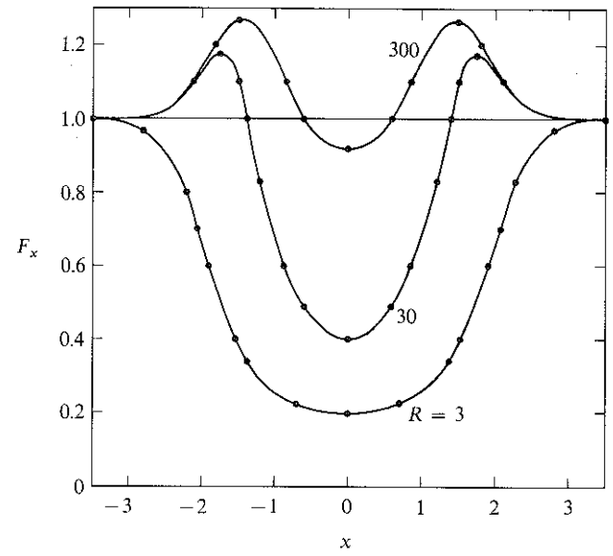

FIGURE 11-14
Emergent flux profiles from spherical atmospheres with $B = 1$, $\varepsilon = 2 \times 10^{-3}$, $T_c = 2$, $T_l = 10^3$, and $\chi \propto r^{-2}$. The source functions in these atmospheres are shown in Figure 14-11.

11-5 Comments on LTE Diagnostics

From the developments presented in the preceeding sections of this chapter we have obtained a very different, and physically a far more satisfying, picture of line-formation from that based on the assumption of LTE. The full multilevel problem must be solved before close *quantitative* agreement with observation can be obtained. It is, nevertheless, worthwhile to consolidate some of the changes in perspective inherent in the new conceptual framework fashioned above, by summarizing here a few of the important differences between the LTE and non-LTE methods of analysis, and by stressing the implications of these differences for the reliability of diagnoses of physical conditions in a stellar atmosphere. An extended discussion of many of these points can be found in (**626**).

In LTE line-formation theory, the source function is linked uniquely to the electron temperature T_e, and the line profile reflects the depth-variation of the Planck function, to within the limits of resolution set by photon diffusion over a single mean-free-path. In contrast, in the non-LTE theory $S_l(\tau)$ is no longer tied to $T_e(\tau)$; instead, it is fixed by the scattering term, which results from the interplay of the rates of photon escape, destruction,

and creation, by all mechanisms. Of these, only collisions couple the creation and destruction events to the local value of T_e. The other photon sources and sinks (which may far overshadow the collisional terms) couple to radiation fields in spectral regions outside of the line. Thus far, only continuum fields have been considered; but, as will be seen in Chapter 12, *all* lines in the entire transition array of the atom are, in principle, involved. To be sure, in the classical theory there was some flexibility implicit in the division of lines into "absorption" and "scattering" lines. But this division was ad hoc, and actually misleading, for we have seen that *all* line source functions have scattering terms, and that the important question is, "What are the sources and sinks that dominate in fixing the level of the scattering term?" The classical theory did not clearly recognize the photoionization category, and was essentially unprepared to cope with the mutual interaction of several lines.

One of the important implications of the above-mentioned changes in the theoretical structure is that, if S_l is no longer uniquely specified by the run of $T_e(\tau)$ (and, in fact, it may be totally unconnected with T_e), then it is clear that we *cannot hope* to infer $T_e(\tau)$ from the emergent intensity $I(\tau = 0, \nu)$. In particular, we have repeatedly found that $S_l(0)$, and hence the central intensity of the line, are almost completely unrelated to the "boundary temperature" T_0, but are determined by transfer over an entire thermalization length, by photons fed into the line wings. Thus the literature containing estimates of stellar boundary temperatures, obtained from line-core intensities using an LTE theory, is little more than a collection of elaborate "maps" of *nonexistent* "territory."

Extending this conclusion, we notice that with an LTE theory, to explain the observed emission cores in the Ca II *H*- and *K*-lines we would have to propose a nonmonotone temperature distribution $T_e^*(z)$, which first decreases outward, then rises, then drops again. Not only would this be inconsistent with the distribution $T_e(z)$ inferred from (LTE!) methods using infrared and continuum data [which show a photospheric drop to a minimum plateau, followed by the chromospheric–coronal rise—cf. Figs. 7-31 and 7-32], but also the particular distribution inferred from the Ca II lines would not yield a fit to other lines (e.g., of Mg II) showing similar reversals. Worse, the use of such a temperature profile would produce emission reversals in *all* lines of sufficiently great strength, even those where none are observed (e.g., $H\alpha$). The entire problem vanishes when we discriminate between the collision-dominated and photoionization-dominated classes; and further, the new approach allows us to understand the varying degree of the effectiveness of the coupling of S_l to T_e for different collision-dominated lines. In short, the non-LTE analysis leads to an enormous improvement in our conceptualization of the physical situation.

One of the primary applications of LTE line-formation theory has been the estimation of stellar abundances. As we have seen, the depth-variation of

$S_l(\tau)$ is often markedly different from $B_\nu(T)$; thus the emergent intensity within the line (and hence its equivalent width) must differ from its LTE value. In a general way, the central depths of collision-dominated lines are increased by departures from LTE, and the equivalent width is increased (which implies that the abundance required to fit an observed width will decrease). For photoionization-dominated lines, non-LTE effects may either increase or decrease the linestrength; a useful summary of expected consequences of departures from LTE on abundances is given in (**225**).

Until recently the question "Do departures from LTE actually lead to significant errors in stellar abundance analyses?" remained open, and was the subject of vigorous and lengthy debate in the literature. As we shall see in §12-4, the question can now be answered in the affirmative in many important cases [see also (**435**)]. This is not to say that LTE abundance estimates are always in error; in many cases they are not. It is, however, apparent that it must be *shown*, and not merely *assumed*, that the application of LTE provides an adequate approximation to reality for any case in question.

Another parameter often inferred from curve-of-growth analyses is the characteristic "microturbulent" velocity in the atmosphere. Leaving aside the questions raised about whether a simple curve of growth adequately describes the complex problem of line-formation in a turbulent medium (discussed in §14-1), it is important to note that the diagnosis of this particular parameter is especially vulnerable to error from non-LTE effects. The position of the flat part of the curve of growth depends sensitively upon the *way*, and the *distance* over which, the line saturates to its thermal value, for these characteristics of the solution determine the *depth* and *width* of the line. For example, as remarked earlier in §11-2, even though coherent scattering may produce a line of the same depth as noncoherent scattering (with the same value of ε), the line will be wider in the latter case, and therefore the equivalent width larger. Calculations, for idealized model atmospheres and two-level atoms, have shown important effects of departures from LTE upon the flat part of the curve of growth (**25**), and recently (cf. §14-1) non-LTE theories that include stochastic velocity fields in the line transfer have been developed. Although there is little doubt that mass motions *do* exist in stellar atmospheres, the accuracy of the actual values assigned will remain in doubt until internally-consistent analytical methods are employed.

We now turn attention to the problem of attempting to match observed stellar spectra, using realistic multilevel model atoms and detailed model atmospheres, in a physically self-consistent solution of the full transfer and statistical equilibrium equations.

12

Non-LTE Line Transfer: The Multilevel Atom

The discussion presented in the previous chapter, based on the use of a highly simplified atomic model in idealized atmospheres, affords deep insight into much of the physics of spectrum line-formation. But to analyze real stellar spectra, it is now necessary (1) to consider much more realistic atomic models, having many levels (perhaps spread over a sequence of ionization stages), giving rise to a multiline transition array; and (2) to solve the combined transfer and statistical equilibrium equations for such atomic models, in fairly elaborate atmospheric models that attempt to describe the physical structure of the star with a high degree of realism. Throughout the discussion we shall assume that we are dealing with an "impurity" species (i.e., other than hydrogen) that has no significant influence on the structure of the atmosphere, and shall therefore regard the atmospheric model as given and *fixed*. It is obvious that with a larger number of levels, the number of transitions and interactions that are possible increases greatly, and both the physical and mathematical nature of the problem become more complex.

Moreover, it is essential to realize that this increased complexity is not merely one of having to treat larger systems of equations, but that there are extremely important new *physical* effects involved. In particular, we shall now find that *the radiation field in any one transition may affect that in every other transition in the atom*, and that in many cases these effects are of overwhelming importance. Furthermore, the way in which the effects manifest themselves is often very subtle, being the result of extremely complicated chains of interactions. This mutual interaction of line radiation fields (and source functions) is referred to as *interlocking*, and the successful treatment of interlocking effects poses the central obstacle to the solution of multilevel transfer problems.

The most straightforward approach to the multilevel problem proceeds from a rather direct extension of the techniques employed in the previous chapter for a two-level atom. Here one writes analytical expressions for the source function in each line in such a way as to isolate explicitly the radiation field in that line, and solves the corresponding transfer equation holding all other terms fixed. This is the *equivalent-two-level atom* approach, in which only one line at a time is considered in the solution of the transfer equation, and the interactions among lines are treated by *iteration*. In the method just described, one has tacitly assumed that the coupling between lines is fairly weak in some sense; but often this is not the case, and the equivalent-two-level-atom approach then becomes unsatisfactory. For example, the individual line components of *multiplets* usually have, as their initial and/or final states, closely-spaced levels that can interact physically, very strongly. In such situations, photons may *switch* from one line to another in the multiplet, and the radiation fields in the lines become strongly interlocked. More generally, in chains of transitions within a complicated transition array, conditions often arise under which photons are rather freely *converted* from one transition, with its corresponding spectral region, to another. In effect, the photon no longer belongs uniquely to a *specific line*, but, to a certain extent, belongs to the *ensemble of radiation fields of the entire set of transitions of the atom*. The profound importance of this point was recognized and emphasized by Jefferies (**334**, Chap. 8; **284**, 177), who advanced the appealing picture that the photons should be considered to be interchangeable members of a *collective pool*. Viewed in this light, it is apparent that it is essential to treat all of the lines and their interactions *simultaneously*, to a high degree of consistency. This may be done quite directly in special cases (e.g., in multiplets with a common lower state). In the general case, strict consistency can be achieved by the *complete linearization method*; this scheme allows fully for all interlocking effects from the outset, and may be considered as the mathematical realization of Jefferies's physical conception of the collective photon pool.

12-1 The Equivalent-Two-Level-Atom Approach

FORMULATION

In our study of the two-level atom, we made use of the statistical equilibrium equations to eliminate analytically the population ratio appearing in the line source function, and thus to obtain an expression of the general form

$$S_l = \left(\int \phi_\nu J_\nu \, d\nu + \alpha\right) \Big/ (1 + \beta) \tag{12-1}$$

where α and β describe the possible ways of creating or destroying a photon. In writing equation (12-1), explicit use is made of the fact that for a *strict* two-level atom, there is only *one* line, hence only one radiation field of relevance. Indeed, we saw that if the atom also has a continuum, then the radiation fields in the two bound–free transitions appear in the terms α and β. Because the line is usually much more opaque than the continua, the latter will often be transparent at the depth of line-formation, and the continuum rates then will be *fixed* (though this is not always the case). If *other* levels and *other* lines are included, it is to be expected that the radiation fields in these transitions will also appear in the terms α and β, and, in addition, that these terms will vary (i.e., cannot be fixed a priori) in the region of line formation. Nevertheless, it is obvious that, however strong the coupling of a given line to other lines may be, the line must always at least respond to the collisional rate coupling the two levels in question, and to the continuum rates from these levels. This suggests that, as a computational strategy, one might attempt to write a mathematical expression for the source function in which the line-scattering, collisional interaction, and continuum radiative and collisional rates from the two levels forming the line appear as *direct rates*, while all other rates are grouped analytically into *net rates*; one hopes in this way to minimize the effect of interlocking at each stage of the calculation.

Consider a line formed between levels l and u. The rate equation for the lower level is

$$\begin{aligned} n_l \left(B_{lu} \int \phi_\nu J_\nu \, d\nu + C_{lu} + \sum_{i<l} A_{li} Z_{li} + \sum_{l<j\neq u} C_{lj} Y_{lj} + R_{l\kappa} + C_{l\kappa}\right) \\ - n_u \left(A_{ul} + B_{ul} \int \phi_\nu J_\nu \, d\nu + C_{ul}\right) \\ = n_l^*(R_{\kappa l} + C_{l\kappa}) + \sum_{l<j\neq u} n_j A_{jl} Z_{jl} + \sum_{i<l} n_i C_{il} Y_{il} \end{aligned} \tag{12-2}$$

and for the upper level we have

$$\begin{aligned} n_u \left(A_{ul} + B_{ul} \int \phi_\nu J_\nu \, d\nu + C_{ul} + \sum_{u>i\neq l} A_{ui} Z_{ui} + \sum_{u<j} C_{uj} Y_{uj} + R_{u\kappa} + C_{u\kappa}\right) \\ - n_l \left(B_{lu} \int \phi_\nu J_\nu \, d\nu + C_{lu}\right) \\ = n_u^*(R_{\kappa u} + C_{u\kappa}) + \sum_{u<j} n_j A_{ju} Z_{ju} + \sum_{u>i\neq l} n_i C_{iu} Y_{iu} \end{aligned} \tag{12-3}$$

In equations (12-2) and (12-3) we have used the *net radiative bracket* Z_{ji} defined by equation (5-58), the *net collisional bracket* Y_{ij} defined by equation (5-71), and the radiative ionization and recombination rates defined by equations (5-66) and (5-67); further n_i^* denotes the LTE population of level i, given by equation (5-14), at the *actual* value of the ion density.

If we solve equations (12-2) and (12-3) for the ratio $(n_l g_u / n_u g_l)$, substitute the result into the relation $S_{lu} = (2h\nu^3/c^2)[(n_l g_u / n_u g_l) - 1]^{-1}$, and make use of the Einstein relations among the transition probabilities, we find

$$S_{lu} = \left[\int \phi_\nu J_\nu \, d\nu + (\varepsilon' + \theta) B_\nu(T_e)\right] \Big/ (1 + \varepsilon' + \eta) \tag{12-4}$$

where ε' is defined as in equation (11-7), while

$$\eta \equiv [a_2 a_3 - (g_l/g_u) a_1 a_4] / [A_{ul}(a_2 + a_4)] \tag{12-5}$$

and

$$\theta \equiv [n_l^* a_1 a_4 (1 - e^{-h\nu/kT})] / [n_u^* A_{ul}(a_2 + a_4)] \tag{12-6}$$

where, in turn,

$$a_1 = R_{l\kappa} + C_{l\kappa} + \sum_{i<l} A_{li} Z_{li} + \sum_{l<j\neq u} C_{lj} Y_{lj} \tag{12-7}$$

$$a_2 = n_l^*(R_{\kappa l} + C_{l\kappa}) + \sum_{l<j\neq u} n_j A_{jl} Z_{jl} + \sum_{i<l} n_i C_{il} Y_{il} \tag{12-8}$$

$$a_3 = R_{u\kappa} + C_{u\kappa} + \sum_{u>i\neq l} A_{ui} Z_{ui} + \sum_{u<j} C_{uj} Y_{uj} \tag{12-9}$$

$$a_4 = n_u^*(R_{\kappa u} + C_{u\kappa}) + \sum_{u<j} n_j A_{ju} Z_{ju} + \sum_{u>i\neq l} n_i C_{iu} Y_{iu} \tag{12-10}$$

Exercise 12-1: Verify equations (12-4)–(12-10).

It is clear that the terms a_1 and a_2 represent, respectively, the loss rate from level l to all levels other than u, and the number of electrons fed into level l from all levels other than u; a_3 and a_4 denote similar quantities for level u.

The resemblance of equation (12-4) to the two-level form is obvious. Note that the populations of *all other levels* appear in the terms a_2 and a_4. This occurs because we have used only *two* of the entire set of rate equations, and hence can eliminate *only* n_l and n_u. It is possible, in principle, to obtain an expression for the source function S_{lu} that does not involve any of the occupation numbers *explicitly*; here the full set of rate equations $\mathscr{A}\mathbf{n} = \mathscr{B}$ is manipulated to give analytical expressions for the quantities α and β, in equation (12-1), in terms of the cofactors of $\mathscr{A}$ [see (**333**; **284**, 187)]. However, the algebra in such a procedure quickly becomes hopelessly unmanageable; and, more important, it is not at all obvious that anything is gained, for the resulting source function then is *nonlinear* in the radiation fields of all other transitions. Moreover, these radiation fields depend upon the occupation numbers in the appropriate levels, hence S_{lu} will still depend *implicitly* upon the n_i's, $(i \neq l, \quad i \neq u)$.

If we write χ_c for the continuum opacity and

$$\chi_{lu} = (B_{lu}h\nu/4\pi)[n_l - (g_l/g_u)n_u]$$

for the line opacity, then the transfer equation at each frequency in the line can be written as

$$\mu(dI_\nu/dz) = -(\chi_c + \chi_l\phi_\nu)I_\nu + \chi_c S_c + \chi_l\phi_\nu S_l \qquad (12\text{-}11)$$

where S_c denotes the continuum source function (which is not necessarily the Planck function). Writing $d\tau_\nu = -(\chi_c + \chi_l\phi_\nu)\,dz$, and substituting for S_l from equation (12-4), equation (12-11) can be rewritten as

$$\mu(dI_\nu/d\tau_\nu) = I_\nu - a_\nu \int \phi_\nu J_\nu\, d\nu - b_\nu \qquad (12\text{-}12)$$

which can be recast into a second-order form, and solved by the standard Feautrier or Rybicki methods described in Chapter 6. A calculation of this type is carried out for *each line* of the entire transition array considered. Note that to compute χ_l, and hence τ_ν, a_ν, and b_ν in equation (12-12), it is necessary to have actual values for n_l and n_u; as these quantities are known only approximately at any stage of the calculation, it is clear that an iteration procedure is required, in which successively more accurate values of the occupation numbers are employed. This iteration may be carried out simultaneously with the more basic iteration required to obtain estimates of the radiative rates in all lines.

In a calculation designed to simulate conditions in a real atmosphere, it is, in general, necessary to account for the variation in the ionization state of the particular atomic species under consideration (perhaps over a sequence of several different ionization stages). Although continua arising from excited states of "impurity" species will generally be transparent (compared to the dominant absorption and emission terms from H and He), so that the radiation field may be considered to be *fixed*, this situation will not usually be true for ground state continua, which normally will be opaque enough to determine their own radiation fields. Thus it is necessary to supplement the line transfer equations (12-12) with corresponding continuum equations, as described in the discussion of the formation of the Lyman continuum in §7-5 [see equations (7-127) through (7-131) and (7-135) through (7-144)]. The complete set of rate equations for a particular impurity species will be composed of as many equations as there are bound levels (over all stages of ionization taken into account), plus a final equation setting the sum of all occupation numbers equal to n_{atom}, the total number density of the species. For a given abundance of the element, relative to hydrogen, n_{atom} is a prespecified function of depth. In overall form these equations are similar to the first M_{He} lines of the matrix displayed, in §5-4, between equations (5-91) and (5-92).

The iteration procedure required to obtain the solution (which consists of the run of the occupation numbers of all the levels with depth) is fairly straightforward in principle, though often complicated in practice. If one starts with an initial estimate of $n_i(z)$ for all atomic levels i, say from LTE relations, then provisional optical depth scales can be constructed. By regarding χ_ν and η_ν as known [via equations (7-1) and (7-2)], an estimate of the radiation field in each transition may be obtained by a *formal solution* of the transfer equation (i.e., S_l is regarded as given). One may use this estimate to compute continuum photoionization and recombination rates. A transfer equation of the form (12-12) for each line (and opaque continuum) is then solved in sequence, supposing, initially, that all the net brackets Z_{ji} and Y_{ij} appearing in equations (12-7) through (12-10), and their continuum analogues, may be set identically to zero. This sequence of calculations produces radiative rates R_{ij} in all lines and opaque continua. The rate equations $\mathscr{A}\mathbf{n} = \mathscr{B}$ are then re-solved, at each depth, to obtain an improved estimate of all the occupation numbers n. All Z_{ji}'s and Y_{ij}'s, and hence η and θ for each line can now be evaluated via equations (12-5) through (12-10) because all of the required rates and level-populations are known. Equations (12-12) are then re-solved, using the new values for τ_ν, η, and θ and one obtains, thereby, new estimates of $\bar{J}_{lu}$ and S_{lu} in every line (and continuum). The rate equations are again re-solved, and the process is iterated to convergence. Once a converged solution is obtained, all the line source functions are known, and line profiles may be calculated for each line.

As mentioned above, the practical implementation of a successful iteration procedure using the equivalent-two-level-atom formalism is often quite complicated, for the rate of convergence (or lack thereof!) may be strongly affected by technical details, such as the way the net rates are computed, source functions are evaluated, and many others. Careful discussions of these points can be found in (**18**, Chap. 4; **23**, 27–63; **23**, 113–132; **187**), as well as in the references cited therein. A more fundamental difficulty with this overall approach is that the iteration procedure may fail to converge but, instead, *stabilizes* on a solution that is *inconsistent*; see the discussion in (**18**, §4.2) and (**23**, 27–63). This failure, which is not too common, and which can sometimes be overcome with special procedures, is, nevertheless, not surprising; as we shall see in the next section, there are many physical situations in which the radiation fields in different lines are very strongly interdependent, in contradiction to the basic assumption of the equivalent-two-level-atom approach. The inconsistency problem is overcome wholly by the complete-linearization method presented in §12.3.

APPLICATION

The equivalent-two-level-atom approach has been extensively applied to calculations of a wide variety of spectra, particularly for the solar atmosphere and for solar-type stars. For example, for the sun, analyses have been made of the Ca II *H*- and *K*-lines (**400**; **401**; **569**), O I and C II lines (**168**), the Mg I *b*-lines and Na I *D*-lines (**21**), and the spectrum of Fe I (**22**), using sophisticated multilevel atoms and elaborate atmospheric models. Very complete spectrum syntheses, using this general approach, of both continuum and line intensities have led to refined photospheric–chromospheric models (**645**; **646**). Further, similar analyses have been made of the Ca II *H*- and *K*-lines in solar-type stars (**56**; **58**).

Space does not permit a discussion of all of these results, and the reader should examine the references cited. It is, nevertheless, of interest to quote a few results (**401**) for the solar Ca II lines, which play a central role in studies of the chromosphere. The *H* and *K* lines arise from the transitions $4s\ ^2S_{\frac{1}{2}} \rightarrow 4p\ ^2P_{\frac{1}{2},\frac{3}{2}}$ (see Figure 12-1); the two upper states are coupled by collisions, and may also decay to the metastable $3d\ ^2D_{\frac{3}{2},\frac{5}{2}}$ states in the infrared triplet $\lambda\lambda$8498, 8542, 8662. As the next levels lie well above the $4p$ level (recall $kT \sim 0.5$ eV in the solar photosphere), the five levels mentioned above, plus the continuum, suffice to provide an accurate description of the physics of line-formation. A solution of the transfer and statistical equilibrium equations for this 5-level Ca^+ atom yields source functions in all 5 lines; the frequency-independent *K*-line source function (obtained assuming complete redistribution) is shown in Figure 12-2, along with the Planck function

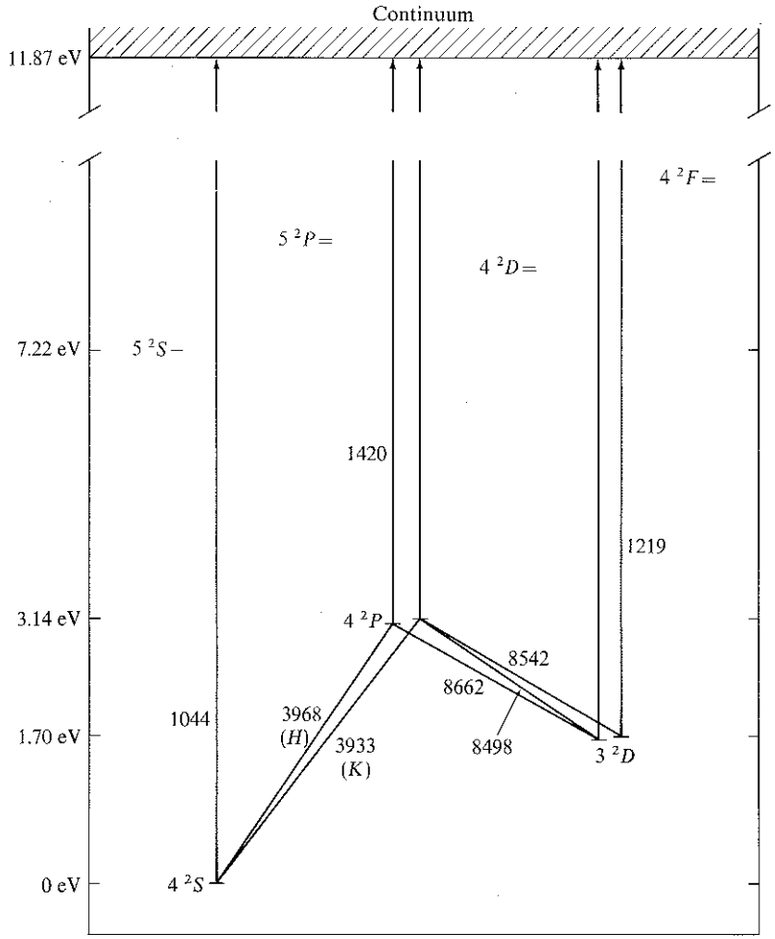

FIGURE 12-1
Energy-level diagram for lower states of Ca^+. The five levels $4s^2S_{\frac{1}{2}}$, $4p^2P_{\frac{1}{2},\frac{3}{2}}$, and $3d^2D_{\frac{3}{2},\frac{5}{2}}$ suffice to describe the line-formation process for the Ca II *H* and *K* lines at $\lambda\lambda$3968 and 3933, and the infrared-triplet $\lambda\lambda$8498, 8542, 8662. Note that the next levels lie fairly high in energy, and hence can be ignored, though interactions with the continuum are included. From (**401**), by permission. (Courtesy of the *Publications* of the Astronomical Society of the Pacific.)

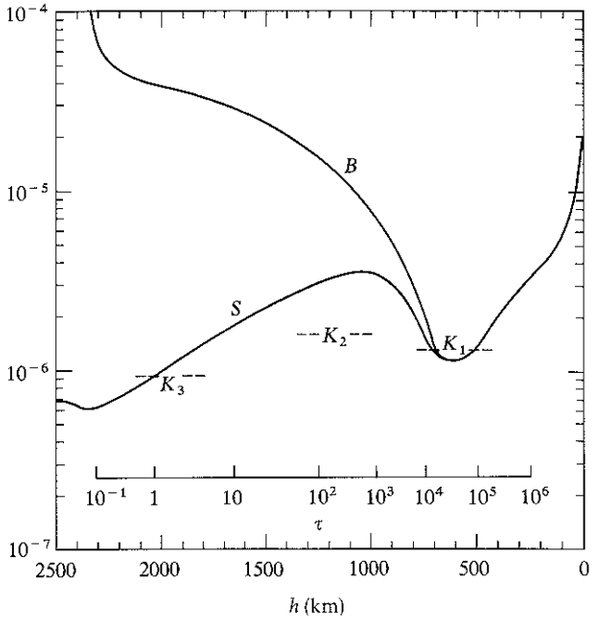

FIGURE 12-2
Ca II K-line source function S, and corresponding Planck-function B, in a model solar atmosphere. Source function is in absolute units: ergs cm^{-2} sec^{-1} hz^{-1} sr^{-1}. Optical depth scale is at line center and dashed lines indicate computed intensities at K_1, K_2, and K_3 (see Fig. 12-3). From (**401**), by permission. (Courtesy of the *Publications* of the Astronomical Society of the Pacific.)

corresponding to the adopted run of temperature $T_e(h)$. The emergent intensity profiles, at $\mu = 1$, for the H- and K-lines and the infrared triplet, are compared with observation in Figures 12-3 and 12-4, respectively. The agreement between theory and observation is quite satisfying. However, when the computed variation of the H- and K-line profiles from center to limb is compared with observation, significant disagreements are found; as we shall see in §13-4, these are removed when the partial coherency of the scattering process, and the resulting frequency-dependence of the source function, are taken into account. Further, the calculations shown in Figures 12-3 and 12-4 represent only the average quiet chromosphere. To match plage profiles (**569**), a different atmospheric model is required; and to match the *detailed* variation of the K-line profile, as observed from point to point on the disk, will require a full treatment of the three-dimensional fine-structure and the velocity fields in the chromosphere. Such calculations have not yet been undertaken in a completely satisfactory fashion, but are needed, both to diagnose details of chromospheric structure, and to answer the question "To what extent

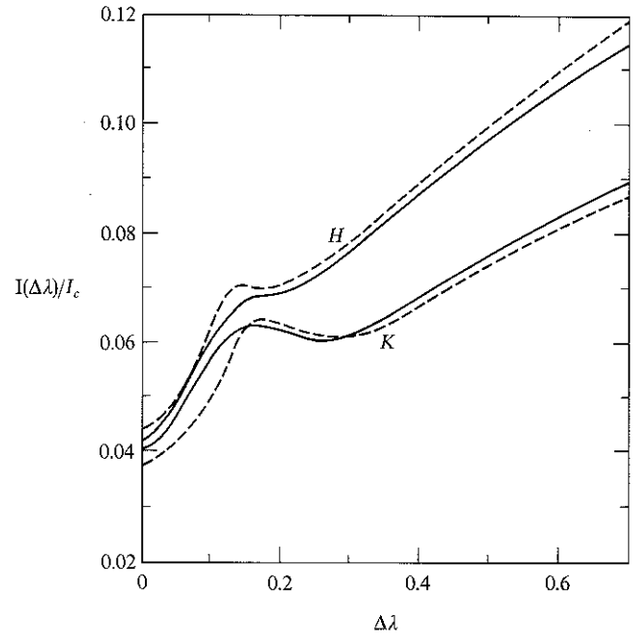

FIGURE 12-3
Computed H- and K-line profiles at $\mu = 1$ (dashed lines) compared with observed profiles (solid lines). From (**401**), by permission. (Courtesy of the *Publications* of the Astronomical Society of the Pacific.)

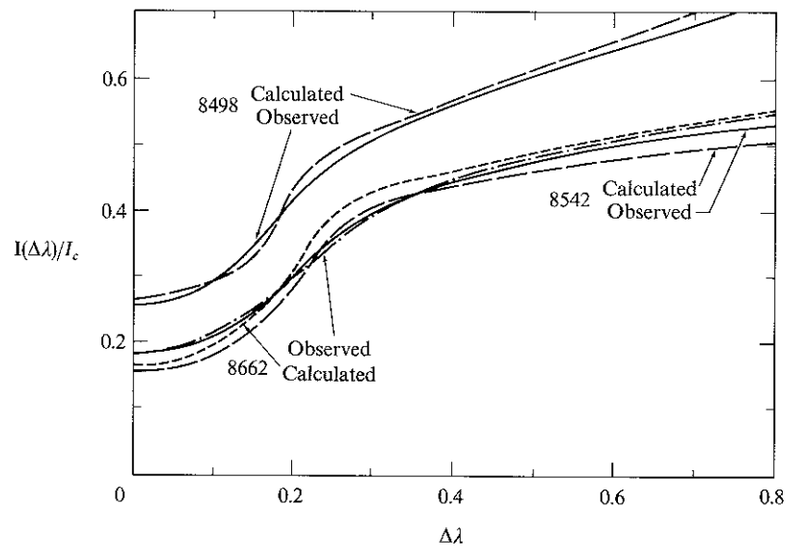

FIGURE 12-4
Computed Ca II infrared-triplet profiles at $\mu = 1$ compared with observed profiles. From (**401**), by permission. (Courtesy of the *Publications* of the Astronomical Society of the Pacific.)

(if any) is it possible to replace the complicated multidimensional structure of the chromosphere with a meaningful horizontally-averaged 'mean chromosphere'?"

12-2 Effects of Level Coupling: Source Function Equality in Multiplets

As was emphasized in the preceding section, the equivalent-two-level-atom approach works best when the coupling terms from two particular levels to all other levels are small, compared to those between the levels themselves and from these levels to the continuum. There are, however, many important physical situations for which this will *not* be the case, and it is important to understand the effects of strong coupling among several levels. These effects have important implications that affect both our perception of line-formation, and our choice of numerical methods used to obtain solutions of multilevel transfer problems. Most of the fundamental physical concepts were outlined by Jefferies (**333**; **334**, Chap. 8), and much of the subsequent theoretical work [e.g., (**567**)] has been based on his ideas. Numerous instructive calculations, using idealized multilevel atomic models and parameterized model atmospheres, have been carried out; see, e.g., (**54**; **18**, Chap. 4; **187**; **213**; **217**; **218**. Rather than attempt to summarize these diverse investigations here, we shall, instead, confine attention to two concrete examples, which illustrate the nature of interlocking effects in a particularly illuminating way.

First we shall consider, briefly, the case of three levels coupled together by two lines of a resonance series, and one subordinate line. The primary result of interest here is *photon degradation* from the higher resonance line into the lower resonance and subordinate lines; this phenomenon is related to fluorescence in nebulae and to the Rosseland cycles discussed in §5-5. Next, we shall devote the remainder of the section to the interlocking of lines within a multiplet; here *photon conversion* may occur as collisions shuffle electrons among the fine-structure levels of the spectroscopic terms involved. As a result, the source functions in the different lines become dependent upon one another, and in a certain limit become *equal* at each depth in the atmosphere. *Source-function equality* in multiplets is of great theoretical importance, for it implies that we may replace several lines within a multiplet by a single representative line, and thereby reduce greatly the amount of computation required. Further, source function equality, when it occurs, allows application of a non-LTE method of analysis of line-profile data, by means of which the physical properties of an atmosphere may be inferred directly with a minimum of theoretical interposition. In keeping with the remainder of the book, in which theoretical prediction rather than analysis of observations has been emphasized, this analytical method will not be discussed here, and the reader is encouraged to refer to the complete discussion by Jefferies (**334**, Chap. 9).

PHOTON DEGRADATION AND CONVERSION

Consider, first, a three-level atom with two resonance lines $1 \leftrightarrow 2$ and $1 \leftrightarrow 3$, and a subordinate line $2 \leftrightarrow 3$. As an example, in hydrogen we would have the $L\alpha$, $L\beta$, and $H\alpha$ lines; in fact, for expository convenience, the three lines will be referred to with these names, even though we may not be dealing with hydrogen. For the present purposes, interactions with the continuum will be ignored, and attention focussed entirely upon the lines. Generally, we can expect the ground-state population n_1 to be vastly larger than the upper-state populations n_2 and n_3; hence the resonance lines will be much more opaque, and will have much larger optical depths, at a given physical depth, than the subordinate lines. Suppose now that all three lines were somehow uncoupled, and were formed in distinct two-level atoms; each would then thermalize at a characteristic depth $\Lambda_{ij} \approx 1/\varepsilon_{ij}$, where ε_{ij} is the collisional destruction probability in transition $i \leftrightarrow j$. Because $h\nu_{23} \ll h\nu_{12}$ (or $h\nu_{13}$) we will generally find $\varepsilon_{23} \gg \varepsilon_{12}$ (or ε_{13}); hence the subordinate line will thermalize in a smaller number of optical depths, measured on its own scale, than the resonance lines.

But this will be overshadowed by the much greater opacity in the resonance lines, and the subordinate-line thermalization will occur at much greater *geometric* depth in the atmosphere. Put another way, the subordinate line is so much more transparent than the resonance lines, that photon-escapes (and hence departures of S from B) first occur in *this* transition, as one proceeds outward from great depth toward the surface of the atmosphere. There will thus be a certain range of depths over which the resonance lines, were they *not* coupled to the subordinate line, would by themselves be thermalized, but within which the subordinate line has become transparent enough to permit escapes. In this critical range, some of the electrons photoexcited by $L\beta$ from level 1 to level 3 will decay into $H\alpha$, and the $H\alpha$ photons will escape; there will thus be a systematic *degradation* of $L\beta$ photons into $H\alpha$ and $L\alpha$ photons, at a rate determined by the branching ratio $A_{32}/(A_{31} + A_{32})$. Note that the inverse process of conversion of $H\alpha$ photons into $L\beta$ is essentially ineffectual. If an $H\alpha$ photon is absorbed, leading to an excitation $2 \rightarrow 3$, *most* of the subsequent emissions from state 3 will be in the $3 \rightarrow 1$ transition; because, however, τ_{13} is so large and ε_{13} is so small, these photons are *trapped*, and merely scatter until a $3 \rightarrow 2$ transition is finally made, and the photon escapes. In this atmospheric region, $L\alpha$ will adjust in such a way as to remain almost in detailed balance (with its source function near to the Planck function), while $L\beta$ excitation is drained, and its source function is depressed.

Ignoring stimulated emissions, we can express "reduced" source functions as $(S_{21}/B_{21}) = (b_2/b_1)$, $(S_{31}/B_{31}) = (b_3/b_1)$, and $(S_{32}/B_{32}) = (b_3/b_2)$, where B_{ij} denotes the Planck function $B(\nu_{ij}, T)$, and $b_i \equiv n_i/n_i^*$. When $L\alpha$ adjusts to detailed balance, $b_2 = b_1$ (recall the discussion pertaining to the ground-state population of H in B-stars given in §7-5), and thus $(S_{32}/B_{32}) = (S_{31}/B_{31})$, so that the "reduced" $L\beta$ and $H\alpha$ source functions become equal below the point of $L\alpha$ thermalization. As we proceed to the outermost layers of the atmosphere, the $L\beta$ line becomes more and more transparent and, ultimately, the probability of *direct* escape exceeds the branching ratio into $H\alpha$, with its consequent escape; at this point the $L\beta$ and $H\alpha$ source functions become uncoupled.

Results from a calculation (**218**) for a three-level atom with the parameters $A_{31}/A_{32} = 10$, $\varepsilon_{12} = \varepsilon_{13} = 10^{-3}$, $\varepsilon_{23} = 9 \times 10^{-2}$, $(\tau_{12}/\tau_{13}) = 6.2$, and $(\tau_{23}/\tau_{13}) = 1.6 \times 10^{-4}$ $[(S_{31}/B_{31})/(S_{32}/B_{32})]$ are shown in Figure 12-5. There we see that $L\alpha$ thermalizes at $\tau_{13} \approx 2 \times 10^2$, which implies $\tau_{12} \approx 10^3$, as expected from ε_{12}. At $\tau_{13} \lesssim 10$, the probability of $L\beta$ escape is comparable to the probability of branching into $H\alpha$; therefore for $\tau_{13} < 10$ the "reduced" $L\beta$ and $H\alpha$ source functions diverge. But below this depth, they rapidly

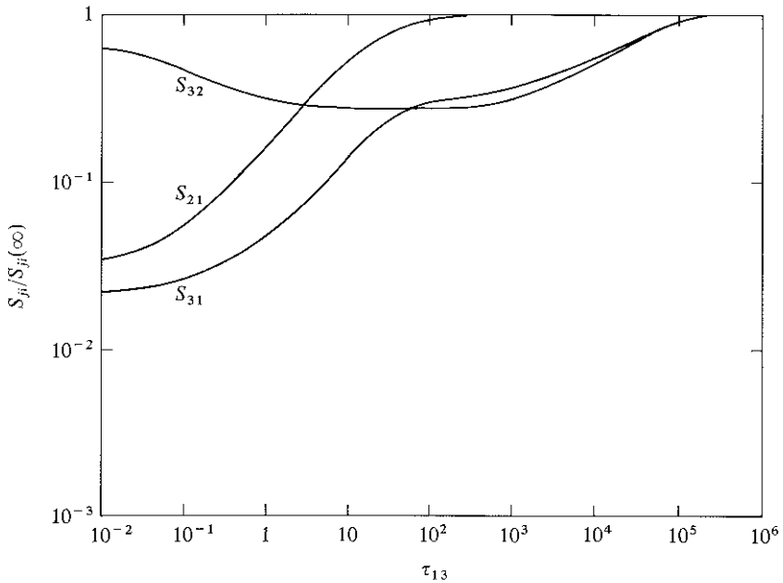

FIGURE 12-5
Source functions for 3-level atom in which both resonance lines, $(1 \leftrightarrow 3)$ and $(1 \leftrightarrow 2)$, and the subordinate line $(2 \leftrightarrow 3)$ are radiatively permitted. *Abscissa*: optical depth in the $1 \rightarrow 3$ line; *ordinate*: "reduced" source functions $[S_{ji}(\tau)/S_{ji}(\infty)]$; $S_{ji}(\infty)$ is equal to the Planck function at the appropriate frequency. From (**218**), by permission.

approach equality, at a value well below that which the $L\beta$ source function would have attained if degradation did not occur. Thermalization in $H\alpha$ (and, along with it, $L\beta$) occurs at about $\tau_{13} \approx 10^5$, which implies $\tau_{23} \sim 10$, as would be expected from ε_{23}. Further discussion and interpretation of the results may be found in the paper cited, but it is clear from what has been said here that there are major effects of interlocking on the source function variations in this problem. See also (**18**, §4.8).

Let us now turn to the case of a three-level atom consisting of a ground state, and two close-lying upper states. We assume that both radiative and collisional transitions can take place between states 1 and 2, and between 1 and 3, but that only collisional transitions occur between states 2 and 3. This model simulates the actual physical situation for a *resonance doublet*, where the upper states are separated by fine-structure splitting. For example, the sodium D-lines arise from transitions from the $3s\ ^2S_{\frac{1}{2}}$ ground state (state 1) to the upper states $3p\ ^2P_{\frac{1}{2}}$ and $3p\ ^2P_{\frac{3}{2}}$ (states 2 and 3).

Consider first the limit in which the $2 \leftrightarrow 3$ collision rate is zero (i.e., the *uncoupled* case). The transitions $1 \leftrightarrow 2$ and $1 \leftrightarrow 3$ can then take place without reference to one another, and the two lines are formed independently. Each line will have a source function that falls below the local Planck function near the surface, and equilibrates to the Planck function at *thermalization depths* given by

$$\Lambda_{12} \approx (A_{21} + C_{21})/C_{21} \tag{12-13a}$$

and

$$\Lambda_{13} \approx (A_{31} + C_{31})/C_{31} \tag{12-13b}$$

Here we have made the simplifying assumptions of Doppler profiles and an absence of strong gradients. In the limit considered here, the run through the atmosphere of the source functions for the two lines will, generally, be different, and at any given depth they will not usually have the same value.

On the other hand, *suppose that collisions occur very rapidly* between levels 2 and 3. In this case, excited electrons are shuffled from one level to the other; we then say that photons have been *converted* back and forth between the two lines. In the transfer problem the two upper states become effectively a single state, and in the extreme limit, the occupation numbers of the two upper states will be proportional to their statistical weights. The source functions for the two lines (which depend on the ratio $n_l g_u/n_u g_l$) then *become equal at each point in the atmosphere.* (Here we ignore the inconsequential differences that may arise because ν_{12} is not *exactly* equal to ν_{13}). In general, the actual situation will be intermediate between the two extremes described above, and we expect that the source functions will be equal only over a limited range in the atmosphere.

Proceeding outward from the deepest layers of the atmosphere, we will find that the source functions in the lines begin to drop below the Planck

function at depths smaller than their thermalization depths; this occurs (cf. §11-2) because the escape probability begins to exceed the destruction probability. But if the collision rates C_{23} and C_{32} are nonzero, there is now a finite chance that the photon will be *switched* from one line to another, rather than be destroyed or escape, and there is now a *conversion length* (**333**; **284**, 177) over which the source functions in the two lines are still coupled together. The lines become uncoupled only when the escape probability exceeds *both* the destruction *and* the conversion probabilities—i.e., when

$$P_e(\tau_{12}) \gtrsim (C_{21} + C_{23})/(A_{21} + C_{21} + C_{23}) \tag{12-14a}$$

and

$$P_e(\tau_{13}) \gtrsim (C_{31} + C_{32})/(A_{31} + C_{31} + C_{32}) \tag{12-14b}$$

From equations (12-14), we see that the two lines could thermalize independently over the entire range predicted by equations (12-13) only if, *everywhere* on that range, $C_{31} > C_{32}$ *and* $C_{21} > C_{23}$ (by a fair margin); if *either* of these conditions is violated, photon conversion plays an essential role. Again assuming Doppler profiles, we can estimate the optical depths from which photons in each line can retain their strict identity, and can emerge, without being either collisionally destroyed *or* converted, as

$$\Lambda^*_{12} \approx (A_{21} + C_{21} + C_{23})/(C_{21} + C_{23}) \tag{12-15a}$$

and

$$\Lambda^*_{13} \approx (A_{31} + C_{31} + C_{32})/(C_{31} + C_{32}) \tag{12-15b}$$

in the $1 \leftrightarrow 2$ and $1 \leftrightarrow 3$ transitions respectively. Let z^*_{12} and z^*_{13} be the *geometric depths* (from the surface) corresponding to Λ^*_{12} and Λ^*_{13}. Then the two groups of photons may propagate independently only over a depth $z^* = \min(z^*_{12}, z^*_{13})$.

From a mathematical point of view, whenever conversion is competetive with thermalization, the usefulness of the iteration procedure employed by the equivalent-two-level-atom approach becomes less clear. It then appears more attractive to consider the two (*or more*) lines, and their effects upon one another, *simultaneously*, and to develop a different computational technique.

OBSERVATIONAL INDICATIONS OF SOURCE FUNCTION EQUALITY

To motivate further the theoretical development, let us consider some of the observational evidence that source function equality in multiplets actually occurs. An excellent example is provided by an extensive set of precise observations of the sodium *D*-lines at various positions on the disk of the Sun (**655**). The emergent intensity at a specified frequency and

disk-position is given by

$$I_\nu(0, \mu) = \mu^{-1} \int_0^\infty [S_c(\tau_c) + \beta_\nu S_l(\tau_c)] \exp\left[-\mu^{-1} \int_0^{\tau_c} (1 + \beta_\nu)\, dt_c\right] d\tau_c \tag{12-16}$$

where τ_c is the continuum optical depth scale, and $\beta_\nu \equiv \chi_l \phi_\nu / \chi_c$. In the core of the line $\beta_\nu \gg 1$, so that, with good accuracy, we may write

$$I_\nu(0, \mu) = \mu^{-1} \int_0^\infty S_l(\tau) \exp(-\tau \phi_\nu / \mu) \phi_\nu \, d\tau \tag{12-17}$$

where τ is now the line optical depth scale. For the *D*-lines one has $(f_{13}/f_{12}) = 2$; thus we may write

$$I_{12}(0, \mu, \nu) = \int_0^\infty S_{12}(\tau_{12}) \exp(-\phi_\nu \tau_{12}/\mu)(\phi_\nu/\mu)\, d\tau_{12} \tag{12-18a}$$

and

$$I_{13}(0, \mu, \nu) = \int_0^\infty S_{13}(\tau_{12}) \exp(-2\phi_\nu \tau_{12}/\mu)(2\phi_\nu/\mu)\, d\tau_{12} \tag{12-18b}$$

Thus clearly $I_{13}(0, \mu, \nu) = I_{12}(0, \frac{1}{2}\mu, \nu)$ *if* S_{12} *and* S_{13} *have a common depth dependence.* In essence, one may compensate for the higher opacity in one line by increasing the path length in the other. When this comparison is carried out, very good agreement between the cores of the two *D*-lines is found, as shown in Figures 12-6a through 12-6c. By way of contrast, Figure 12-6d shows the inverse comparison of $I_{13}(0, \frac{1}{2}\mu, \nu)$ with $I_{12}(0, \mu, \nu)$, and demonstrates the genuine significance of the agreement shown in the other figures. The disagreement in the wings arises from the increasingly large contributions from the continuum, which invalidate the assumptions required to write equation (12-17).

These striking observational results provide an impetus to examine in detail the conditions under which source function equality can occur (aside from the trivial case of LTE), and to develop methods well suited to handle the transfer problem for multiplets. The strong agreement shown in Figure 12-6 is so impressive that, in some of the early analyses (**655**; **656**), it was concluded that there must be *strict* source function equality in the two lines all the way to the surface. To obtain such equality, one would have to impose the exacting requirement that $C_{23} \gg A_{21}$ and $C_{32} \gg A_{31}$; these conditions cannot actually be met by the collisional rates for the *D*-lines in the solar atmosphere, and an apparent contradiction arises. Subsequent work (**16**; **51**) has shown, however, that the requirements stated above are far too stringent, and that one needs to have only $(C_{32}/A_{31}) > (C_{31}/A_{31})$ and $(C_{23}/A_{21}) > (C_{21}/A_{21})$ (conditions that *are* met for the solar *D*-lines) to obtain a very near (though not *exact*) equality of the source functions. These conditions are sufficient to produce profiles that look identical to within the accuracy of the observations.

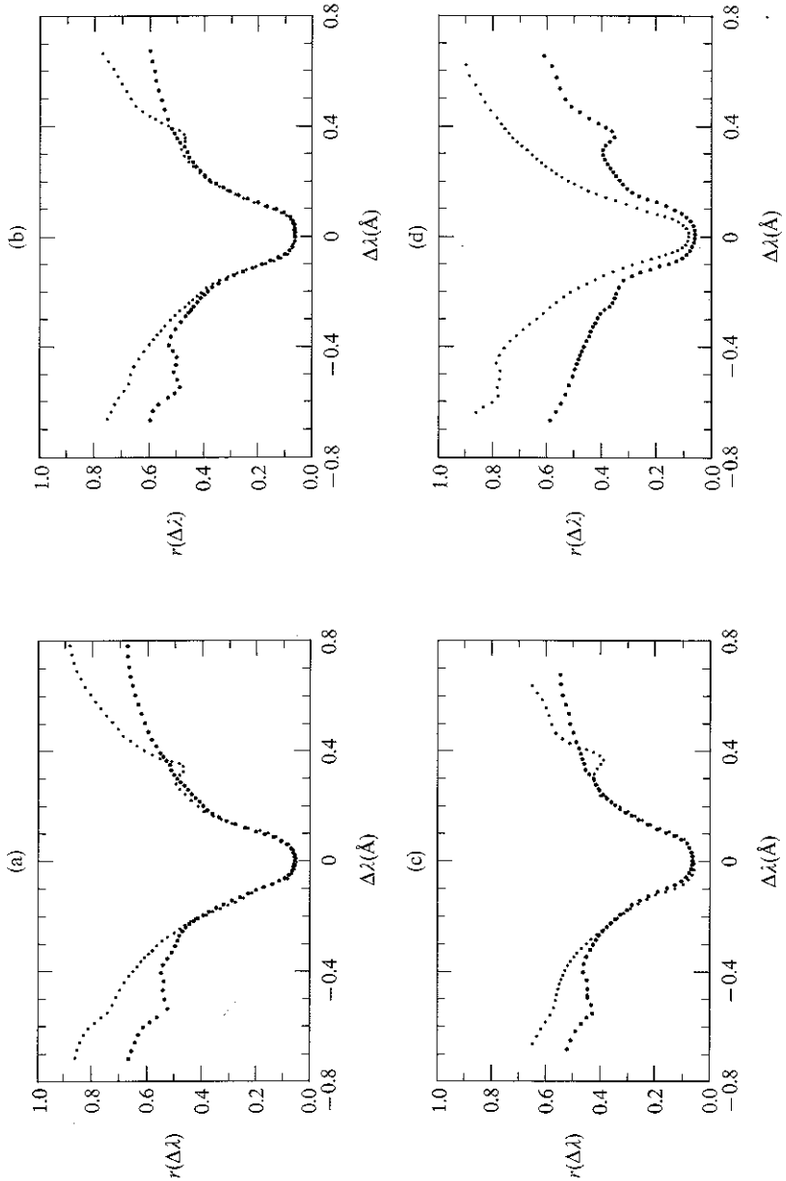

FIGURE 12-6
Comparison of $I_{12}(0, \frac{1}{2}\mu, \nu)$ with $I_{13}(0, \mu, \nu)$ for solar sodium *D*-lines: (a) for $\mu = 1.0$; (b) for $\mu = 0.8$; and (c) for $\mu = 0.6$. (d) Inverse comparison of $I_{12}(0, \mu, \nu)$ with $I_{13}(0, \frac{1}{2}\mu, \nu)$ for $\mu = 1$; note complete lack of agreement in this case. From **(655)**, by permission.

SOLUTION OF THE TRANSFER EQUATION IN MULTIPLETS

Let us now consider how the coupled transfer problems for lines in a multiplet, arising from a single lower level to two (or more) upper levels that interact via collisions, are formulated and solved. For simplicity, stimulated emissions and coupling to the continuum will be neglected, and the differences in line frequencies ignored; these may, of course, be taken into account, but the algebra becomes much more complicated without a compensating enhancement of the physical content. The source function in the $(1 \rightarrow j)$ transition is

$$S_{1j} = (2h\nu^3/c^2)(g_1 n_j / g_j n_1) \tag{12-19}$$

The required population ratios are obtained from rate equations; for the case of only two upper levels these are:

$$n_2(A_{21} + C_{21} + C_{23}) = n_1(B_{12}\bar{J}_{12} + C_{12}) + n_3 C_{32} \tag{12-20a}$$

and

$$n_3(A_{31} + C_{31} + C_{32}) = n_1(B_{13}\bar{J}_{13} + C_{13}) + n_2 C_{23} \tag{12-20b}$$

where

$$\bar{J}_{1j}(\tau) \equiv \int_{-\infty}^{\infty} J_{1j}(\tau, \nu)\phi_j(\tau, \nu)\, d\nu = \frac{1}{2}\int_{-\infty}^{\infty} d\nu\, \phi_j(\tau, \nu) \int_{-1}^{1} d\mu\, I_{1j}(\tau, \mu, \nu) \tag{12-21}$$

Making use of the Einstein relations, and detailed-balancing relations of the form $C_{ij} = (n_j/n_i)^* C_{ji}$, and defining $\varepsilon_{12} \equiv C_{21}/A_{21}$, $\eta_{12} \equiv C_{23}/A_{21}$, $\varepsilon_{13} \equiv C_{31}/A_{31}$, and $\eta_{13} \equiv C_{32}/A_{31}$, we may rewrite equations (12-20), in terms of source functions of the form of (12-19), as

$$S_{12} = (\bar{J}_{12} + \varepsilon_{12}B_\nu + \eta_{12}S_{13})/(1 + \varepsilon_{12} + \eta_{12}) \tag{12-22a}$$

and

$$S_{13} = (\bar{J}_{13} + \varepsilon_{13}B_\nu + \eta_{13}S_{12})/(1 + \varepsilon_{13} + \eta_{13}) \tag{12-22b}$$

These expressions display transparently the *linear dependence* between S_{12} and S_{13}. We see that if $\eta_{12} = \eta_{13} \equiv 0$, the source functions are *uncoupled*. But if either $\eta_{12} \gg 1$ or $\eta_{13} \gg 1$, then $S_{12} \rightarrow S_{13}$. We see further, that even if the strong inequalities just mentioned are not satisfied, merely having $\eta_{1j} \gg \varepsilon_{1j}$ and $\eta_{1j}S_{1k} \gg \varepsilon_{1j}B_\nu$ will imply a near-equality of S_{12} and S_{13}, for then the physical source-sink terms in each source function will be dominated by the *other* line. The η_{1j} terms are thus expected to influence the solution in a way reminiscent of the collisional versus photoionization domination found in the two-level case (i.e., the lines in a multiplet may be "conversion-dominated").

In view of the linear dependence discussed above, it is evident that a simultaneous solution for both source functions is mandatory. Let us

consider a single ground state and an arbitrary number of upper levels. Equations (12-22) can be replaced by the general form

$$S_{1j} = \alpha_j \bar{J}_{1j} + \sum_k \beta_{jk} S_{1k} + \gamma_j \tag{12-23}$$

where the sum extends over all upper levels. This system can be solved (at each depth-point in the medium) to obtain an expression of the form

$$S_{1j}(\tau) = \sum_k a_{jk}(\tau) \bar{J}_{1k}(\tau) + b_j(\tau) \tag{12-24}$$

Exercise 12-2: Write explicit analytical expressions, of the form of equation (12-24), for the case of only two upper levels, starting from equations (12-22).

To write the transfer equation we now adopt the optical depth τ_{1i} in some particular line $(1 \rightarrow i)$ as standard, and write $\gamma_j \equiv (d\tau_{1j}/d\tau_{1i}) = (f_{1j}/f_{1i})$. Then for each line $(1 \rightarrow j)$ we have

$$\mu[dI_{1j}(\tau, \mu, \nu)/d\tau] = \gamma_j \phi_j(\tau, \nu)[I_{1j}(\tau, \mu, \nu) - S_{1j}(\tau)] \tag{12-25}$$

We introduce depth, angle, and frequency discretizations $\{\tau_d\}$, $\{\mu_m\}$, and $\{\nu_n\}$, and appropriate quadrature sums required to perform the integrations indicated in equation (12-21). Then defining

$$u_{j,dmn} \equiv u_{1j}(\tau_d, \mu_m, \nu_n) = \frac{1}{2}[I_{1j}(\tau_d, +\mu_m, \nu_n) + I_{1j}(\tau_d, -\mu_m, \nu_n)] \tag{12-26}$$

and rewriting equation (12-21) as

$$\bar{J}_{j,d} \equiv \bar{J}_{1j}(\tau_d) = \sum_n \sum_m w_{mn} \phi_j(\tau_d, \nu_n) u_{j,dmn} \tag{12-27}$$

the transfer equation can be written in the discrete form

$$\mu_m^2 (d^2 u_{j,dmn}/d\tau^2) = (\gamma_{j,d} \phi_{j,dn})^2 \left(u_{j,dmn} - \sum_{k=2}^{L+1} a_{jk,d} \bar{J}_{j,d} - b_{j,d} \right) \tag{12-28}$$

Here $\phi_{j,dn} \equiv \phi_j(\tau_d, \nu_n)$, $a_{jk,d} \equiv a_{jk}(\tau_d)$, $b_{j,d} \equiv b_j(\tau_d)$,

and $\gamma_{j,d} \equiv \gamma_j(\tau_d)$

the sum extends over all L lines (or upper levels) considered. An equation of the form of (12-28) can be written for each line.

To solve the system we may use the Rybicki method. Define

$$\mathbf{u}_{j,mn} \equiv (u_{j,1mn}, \ldots, u_{j,dmn}, \ldots, u_{j,Dmn}) \tag{12-29}$$

and

$$\bar{\mathbf{J}}_j \equiv (\bar{J}_{j,1}, \ldots, \bar{J}_{j,d}, \ldots, \bar{J}_{j,D})^T \tag{12-30}$$

which describe the depth-variations of these quantities on the range $\tau_1 \leqslant \tau \leqslant \tau_D$. Then the system (12-28) can be written in the general form

$$\mathbf{T}_{j,mn} \mathbf{u}_{j,mn} + \sum_{k=2}^{L+1} \mathbf{U}_{k,mn} \bar{\mathbf{J}}_k = \mathbf{K}_{j,mn} \tag{12-31}$$

for each $(j = 2, \ldots, L+1)$, $(m = 1, \ldots, M)$, $(n = 1, \ldots, N)$. Here $\mathbf{T}$ is a $(D \times D)$ tridiagonal matrix, the $\mathbf{U}$'s are $(D \times D)$ diagonal matrices, and $\mathbf{K}$ is a vector of dimension D. These systems are solved in succession to yield, in effect,

$$\mathbf{u}_{j,mn} = \sum_{k=2}^{L+1} \mathbf{C}_{jk,mn} \bar{\mathbf{J}}_k + \mathbf{D}_{j,mn} \tag{12-32}$$

where $\mathbf{C}$ is a full $(D \times D)$ matrix. Equations (12-32) are substituted into the matrix representations of equation (12-27), namely $\bar{\mathbf{J}}_j = \sum_{mn} \mathbf{W}_{j,mn} \mathbf{u}_{j,mn}$, to generate a final system of the form

$$\sum_{k=2}^{L+1} \mathbf{P}_{jk} \bar{\mathbf{J}}_j = \mathbf{Q}_j, \quad (j = 2, \ldots, L+1) \tag{12-33}$$

Here each $\mathbf{P}_{jk}$ is a $(D \times D)$ matrix, so the whole system is of order $LD \times LD$. The system is then solved by standard numerical methods, and yields the full depth-variation in all lines simultaneously. When stimulated emission terms are included, the system becomes nonlinear and requires an iterative solution. A linearization method for the problem is described in (**23**, 1); we shall not discuss this particular procedure, but instead will describe, in §12-3, a more general method that handles this problem and many others from the outset.

The sodium D-lines have been studied (**51**), using schematic atomic and atmospheric models, with an integral-equation method essentially equivalent to the system described above. For these particular lines the following relations among atomic parameters exist: $g_1 = 2$, $g_2 = 2$, $g_3 = 4$, $A_{31} = A_{21}$, $B_{13} = 2B_{12}$, $C_{21} = C_{31}$; hence $\varepsilon_{12} = \varepsilon_{13}$ and $\eta_{12} = 2\eta_{13}$. Adopt the $1 \leftrightarrow 3$ transition as standard so that $\gamma_2 = \frac{1}{2}$ and $\gamma_3 = 1$. Solutions were obtained for a range of typical values of the parameters ε_{13} and η_{13}, and depth- variations of the Planck function. Results for cases with $B_\nu \equiv 1$, $\varepsilon = 10^{-4}$, and $\eta = \eta_{13} = 0$, 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}, and 1, are shown in Figure 12-7. We see that, while η_{13} would have to be larger than 1 to guarantee strict source function equality to the very surface, much smaller values of η_{13} yield equality for $\tau \gtrsim 1$, which is the significant range for determining the emergent intensity. The emergent specific intensity for $\mu = 1.0$, 0.8, and 0.6 is shown in Figure 12-8. Clearly the profiles agree very closely when $\eta = 10^{-3}$, and are indistinguishable when $\eta = 10^{-2}$. These calculations show that, for all practical purposes, source function equality *can* occur for $\eta \ll 1$; indeed this result is not surprising, for we would, in fact, *expect*

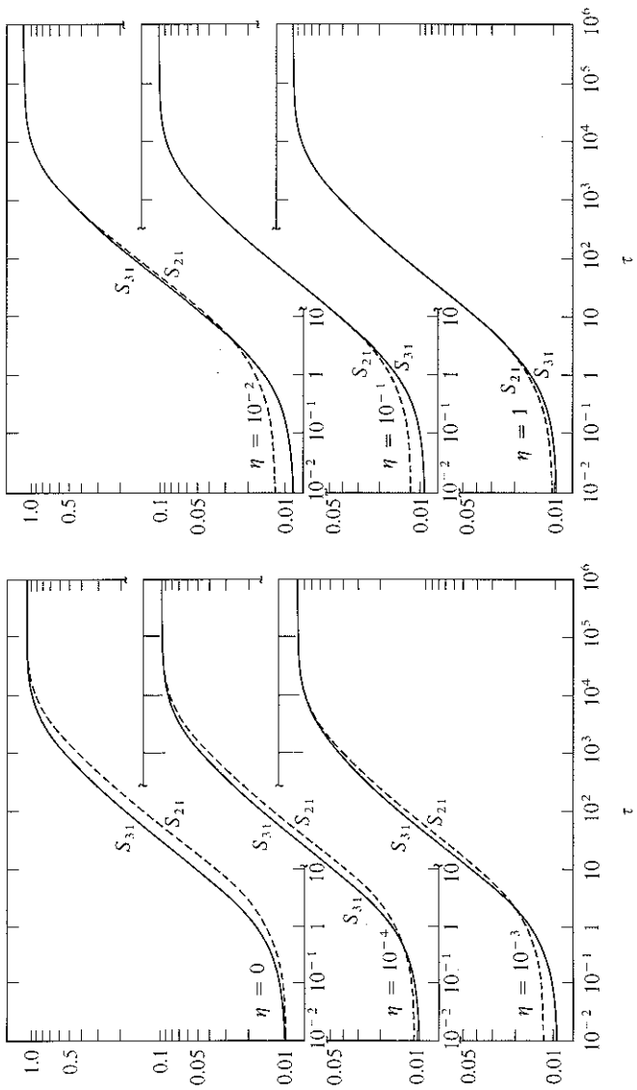

FIGURE 12-7
Source functions for an idealized sodium D-line multiplet in a semi-finite atmosphere, assuming $B_\nu = 1$, $\varepsilon = 10^{-4}$ and various values of $\eta = \eta_{13}$. From (**51**), by permission.

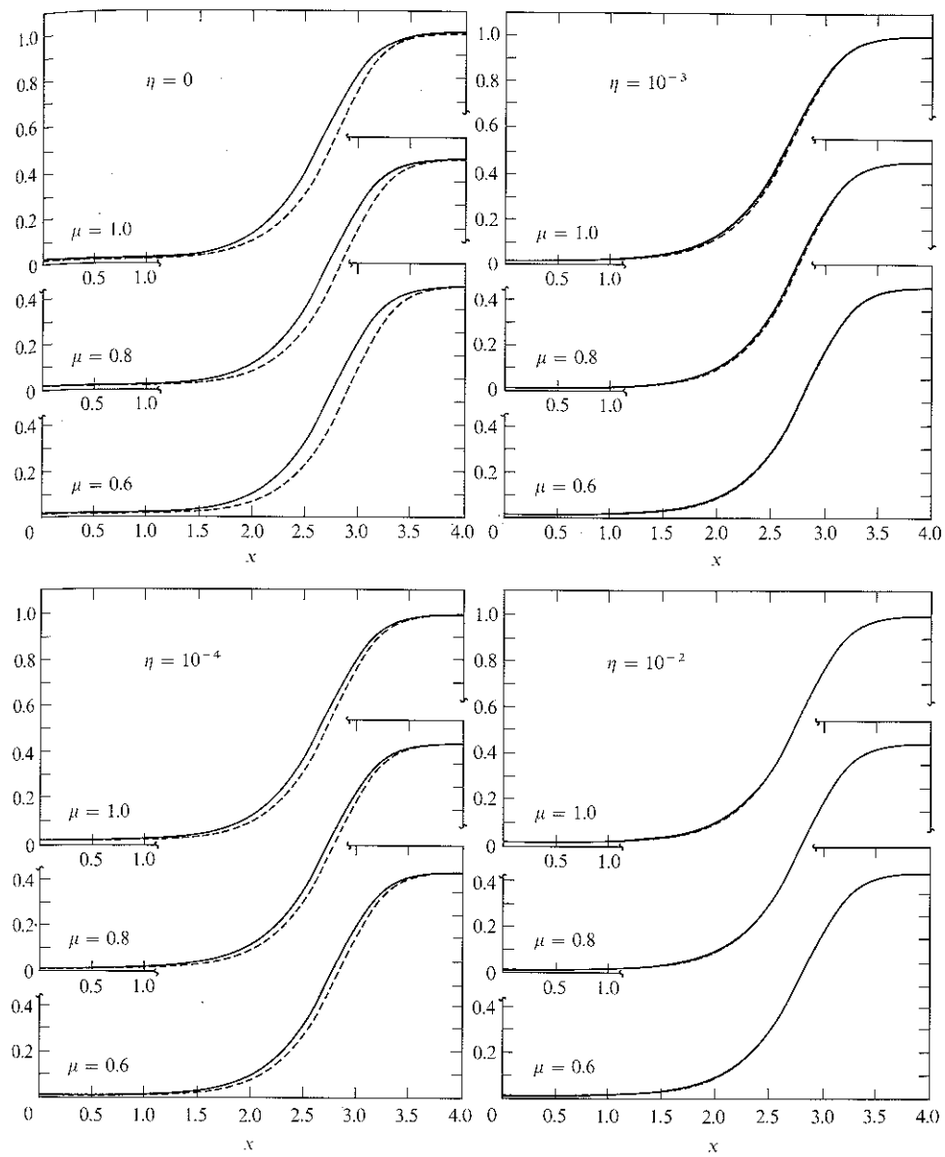

FIGURE 12-8
Line profiles comparing $I_{12}(0, \frac{1}{2}\mu, \nu)$ with $I_{13}(0, \mu, \nu)$ for various values of η, at $\mu = 1$, 0.8, and 0.6. Solid curves give I_{13}, dashed curves give I_{12}. *Abscissa*: displacement from line center in units of Doppler widths. From (**51**), by permission.

conversion to dominate individual line-thermalization as soon as $\eta \gg \varepsilon$. The effects of a schematic temperature gradient were also studied, and it was found that, when $B_\nu(\tau)$ increases inward, the value of η required for source function equality decreases (compared to the $B_\nu \equiv 1$ case), and increases when $B_\nu(\tau)$ decreases inward. These results, however, may depend upon the specific forms chosen for $B_\nu(\tau)$, and may not be general.

12-3 The Complete Linearization Method

In the light of the discussion presented thus far in this chapter, it can be appreciated that the effects of interlocking in the multilevel line-formation problem are important, complex, and lead to systems of equations that may be ill-conditioned. Neither of the methods discussed thus far is entirely satisfactory for multilevel problems. The equivalent-two-level-atom formulation is ill-posed when there is strong coupling among lines. The method developed to treat multiplets is too specialized. What is needed is a method that is general, flexible, and computationally robust, and that handles the physical complexities of the problem. These requirements are met by the *complete linearization method* (**42**). In this method, the *full* set of rate equations is incorporated into the transfer equation, from the outset, by linearizing in terms of all occupation numbers and their dependences upon the radiation field. The rate equations themselves constitute the physical prescription of how the photons interact with the material, are created, destroyed, converted, degraded, and interlock within the collective photon pool envisioned by Jefferies. The transfer equations determine how this information is propagated from one depth to another. By means of the linearization procedure, a set of equations is developed that describes fully consistently (to first order) the response of the material, at *every* point in the atmosphere, to the radiation field at *any* frequency and *any* depth-point and, reciprocally, the response of the radiation field at *all* frequencies and *all* points in the medium to a change in material properties at *any* point. When this system is solved (iteratively), one has a simultaneous solution of the full set of rate equations and transfer equations that reflects both the global nature of the transfer problem (inherently large destruction/conversion lengths interplaying with an open boundary surface) and the intricate infrastructure of the statistical equilibrium equations. In practice, this method has provided a means of treating extremely complex model atoms with a very high degree of realism.

We shall examine here the statistical equilibrium problem for a multilevel "impurity" atom that has no effect upon the structure of the atmosphere. We regard the model atmosphere as *given*—i.e., the temperature (T), electron density (n_e), mass density (ρ), and the total number density of the species under consideration (n_{atom}) are all specified, fixed, functions of depth. The atom is assumed to have L discrete levels, distributed, perhaps, over several

ionization stages. The rate equations are formulated to describe all interactions among all levels, and to fix the total number density at the specified value n_{atom}. These equations will be of the general form of the first M_{He} lines of the matrix displayed in §5-4, supplemented by an equation of the form $\sum_{(i,\,j)} n_{ij} = n_{\text{atom}}$. For ease of exposition we make the simplifying assumption that, at any chosen frequency ν_n, there is only *one* transition (say $l \leftrightarrow u$) of the atom that can interact (bound–bound or bound–free) with the radiation field, and that all overlapping sources of opacity and emission are fixed.

The transfer equation to be solved at each frequency ν is

$$\partial^2(f_\nu J_\nu)/\partial\tau_\nu{}^2 = J_\nu - (\eta_\nu/\chi_\nu) \tag{12-34}$$

From equations (7-1) and (7-2) we can write the opacity and emissivity in a bound–bound transition as

$$\chi_\nu = \alpha_{lu}(\nu)[n_l - (g_l/g_u)n_u] + X_\nu \tag{12-35a}$$

$$\eta_\nu = (2h\nu^3/c^2)\alpha_{lu}(\nu)(g_l/g_u)n_u + E_\nu \tag{12-35b}$$

and that in a bound–free transition as

$$\chi_\nu = \alpha_{lu}(\nu)[n_l - n_u(n_l/n_u)^* e^{-h\nu/kT}] + X_\nu \tag{12-36a}$$

$$\eta_\nu = (2h\nu^3/c^2)\alpha_{lu}(\nu)n_u(n_l/n_u)^* e^{-h\nu/kT} + E_\nu \tag{12-36b}$$

where X_ν and E_ν are fixed. To unify the notation, write

$$\chi_\nu = \alpha_{lu}(\nu)[n_l - G_{lu}(\nu)n_u] + X_\nu \tag{12-37a}$$

$$\eta_\nu = (2h\nu^3/c^2)\alpha_{lu}(\nu)G_{lu}(\nu)n_u + E_\nu \tag{12-37b}$$

where $G_{lu}(\nu) \equiv (g_l/g_u)$ for bound–bound transitions, and $G_{lu}(\nu) \equiv n_e\Phi_l(T)\exp(-h\nu/kT)$ for bound–free transitions; here $\Phi_l(T)$ is the Saha–Boltzmann factor for level l [cf. equation (5-14)]. The system is discretized in depth $\{m_d\}, (d = 1, \ldots, D)$, and frequency $\{\nu_k\}, (k = 1, \ldots, K)$. We then can write a transfer equation of the form of equation (7-37) at each frequency ν_k; these equations are linearized to obtain an equation of the form of (7-39), where now $\delta\chi_{dk}$ and $\delta\eta_{dk}$ can be expressed simply as

$$\delta\chi_{dk} = \alpha_{lu}(\nu_k)[\delta n_{l,\,d} - G_{lu}(\nu_k)\,\delta n_{u,\,d}] \tag{12-38a}$$

$$\delta\eta_{dk} = (2h\nu_k{}^3/c^2)\alpha_{lu}(\nu_k)G_{lu}(\nu_k)\,\delta n_{u,\,d} \tag{12-38b}$$

The δn's can be expressed in terms of changes in the radiation field only, as T and n_e are assumed fixed, and from equation (5-102) or (7-157) we can write

$$\boldsymbol{\delta n}_d = \sum_{k=1}^{K} \left(\frac{\partial \mathbf{n}}{\partial J_k}\right)_d \delta J_{dk} \tag{12-39}$$

where *analytical* expressions for $(\partial \mathbf{n}/\partial J_k)$ are given by equation (5-108). By means of equations (12-38) and (12-39), it is now easy to eliminate all the δn's in the linearized transfer equation *analytically*, in terms of the δJ's only (**33**), and to obtain equations in the standard Feautrier form

$$-\mathbf{A}_d\,\delta\mathbf{J}_{d-1} + \mathbf{B}_d\,\delta\mathbf{J}_d - \mathbf{C}_d\,\delta\mathbf{J}_{d+1} = \mathbf{L}_d \tag{12-40}$$

where

$$\delta\mathbf{J}_d \equiv (\delta J_{d1}, \ldots, \delta J_{dk}, \ldots, \delta J_{dK})^T \tag{12-41}$$

and $\mathbf{A}_1 = \mathbf{C}_D \equiv 0$.

Exercise 12-3: Write explicit expressions for the elements of the matrices $\mathbf{A}_d$, $\mathbf{B}_d$, and $\mathbf{C}_d$, and the vector $\mathbf{L}_d$ in equation (12-40).

In the formulation of the linearization procedure described in §7-5, both the δJ's and δn's appeared explicitly. By eliminating the δn's the system size has been reduced from $[(K + L) \times (K + L)]$ to $(K \times K)$, which saves computer storage; *but* now **A** and **C** are *full*, whereas in the earlier formulation [cf. equation (7-159)] both **A** and **C** were *diagonal* (omitting the constraint of hydrostatic equilibrium which is not relevant in the present context). All of the operations in the elimination scheme given by equations (6-40) and (6-41) now involve full matrices, and the computing time therefore increases; the overall scaling is $T = c\,DK^3$, compared to the older scaling $T' = c'\,D(K + L)^3$, but c is much larger than c'.

The solution of equations (12-40) is carried out iteratively, as described below. The great power of the method is that equations (12-40) account for the full depth-to-depth and frequency-to-frequency coupling in the problem. That is, the effect of δJ_{dk} at (τ_d, ν_k) upon $J_{d'k'}$ at all other $(\tau_{d'}, \nu_{k'})$ is given consistently, and thus photon propagation within the collective pool is taken fully into account.

In actual application of the method, we must now consider how to obtain a starting solution, and how to treat a large number of lines, as required by a realistic model atom. These questions are related, and the latter is of particular importance. For example, suppose the model atom has, say, 20 levels. On combinatorial grounds there could be of the order of 200 transitions; if we require, say, 10 frequencies per line profile, the problem becomes unmanageable. But of course spectroscopic selection rules severely limit the number of transitions that are actually possible in the spectrum, and among 20 levels one would typically have about 30 permitted lines. Of these, only a particular subset will directly influence those levels of primary interest in the treatment of definite limited set of lines in the spectrum, and usually it is possible to divide the transition array into a set of "primary" lines for which a strictly self-consistent treatment is necessary and a set of "secondary" lines for which a less precise treatment is adequate. This division is exploited in the initialization procedure.

To initialize the problem, the model atmosphere is used to determine background opacities and emissivities, and the radiation field is computed, assuming LTE populations for the atom under study. This radiation field is used to calculate photoionization rates, and the statistical equilibrium equations are solved assuming radiative detailed balance in the lines, but not in the continua; in this way one obtains the correct asymptotic solution at depth. It is then possible to do an equivalent-two-level-atom solution to find $\bar{J}_{lu}$ in each line.

The resulting $\bar{J}$'s allow the radiative rates in the lines to be computed, and the complete statistical equilibrium equations can then be solved to obtain improved level-populations. This process may be iterated, and may be extended to include the analogue of the equivalent-two-level-atom solution for the opaque continua as described in §7-5 [cf. equations (7-135) through (7-144)]. The process outlined here could, in principle, be iterated to consistency, and one would then have the method of §12-1; however, in the present context, only one or two iterations are performed, and the results are adopted as initial estimates of the n's, J_ν's, and Eddington factors required to start the linearization procedure. At this point, and henceforth, the radiative rates in the "secondary" lines are held fixed at their equivalent-two-level-atom values, and only the "primary" lines are treated explicitly in the linearization. This will provide an adequate treatment of the *secondary* lines *if* they have been chosen astutely, *and* if they (a) are very weak, and therefore dominated by overlapping absorption–emission processes, or (b) are, in fact, "isolated", and therefore accurately described by a two-level atom, or (c) are only *very weakly* coupled to the lines of primary interest, even though they are members of, say, a multiplet, and the two-level-atom approximation is poor.

Starting from the initial solution just described, the linearized equations (12-40) are now solved for $\delta\mathbf{J}$ at all depths. The resulting $\delta\mathbf{J}$'s are used to obtain a more accurate estimate of the radiation field; the radiative rates are then updated with this revised field, and the statistical equilibrium equations are re-solved for new level populations. With the new occupation numbers, opacities and emissivities may be computed, and a formal solution (Λ-iteration) is then performed to obtain the radiation field and Eddington factors (this step also provides a smoothing of the solution). The convergence of the method is ordinarily very swift, yielding $\|\delta J_\nu/J_\nu\| \lesssim 10^{-5}$ in 4 or 5 iterations, typically with order-of-magnitude improvement between successive iterations. This scheme is the one used to carry out most of the work discussed in §12-4. A detailed description of a particular version of the method, including a program adapted for calculations in the solar atmosphere, is given in (**36**).

A difficulty with the Feautrier-type approach developed above is that it becomes costly for computations involving large numbers of frequencies.

The problem can be overcome, however, with a new method (**37**) that uses a Rybicki-type elimination. Again we assume that at frequency ν_k there is only one transition (t), arising from levels l and u. At this particular frequency, the linearized transfer equation (7-39) can be written, in view of equations (12-38), as

$$\mathbf{T}_k\,\boldsymbol{\delta}\mathbf{J}_k + \mathbf{L}_k\,\boldsymbol{\delta}\mathbf{n}_l + \mathbf{U}_k\,\boldsymbol{\delta}\mathbf{n}_u = \mathbf{R}_k \tag{12-42}$$

where each of the $\boldsymbol{\delta}$-vectors contains the *depth* variation of the quantity—e.g.,

$$\boldsymbol{\delta}\mathbf{J}_k = (\delta J_{1k}, \ldots, \delta J_{dk}, \ldots, \delta J_{Dk})^T \tag{12-43}$$

and the matrices $\mathbf{T}_k$, $\mathbf{L}_k$, and $\mathbf{U}_k$ are of dimension $(D \times D)$ and are tridiagonal. Equation (12-42) may be solved to obtain an equation of the form

$$\boldsymbol{\delta}\mathbf{J}_k + \mathscr{L}_k\,\boldsymbol{\delta}\mathbf{n}_l + \mathscr{U}_k\,\boldsymbol{\delta}\mathbf{n}_u = \mathscr{R}_k \tag{12-44}$$

where $\mathscr{L}$ and $\mathscr{U}$ are now full matrices. Now the basic radiation-dependent quantities entering the rate equations are radiative *rates* integrated over the transitions in question. We therefore introduce the variations in the net rates, defined as

$$\begin{aligned}(\boldsymbol{\delta}\mathbf{Z}_t)_d &= n_{l,d}\,\delta R_{lu,d} - n_{u,d}\,\delta R_{ul,d} \\ &= \sum_k \left[4\pi w_k \alpha_{lu}(\nu_k)/h\nu_k\right]\left[n_{l,d} - G_{lu}(\nu_k) n_{u,d}\right]\delta J_{dk}\end{aligned} \tag{12-45}$$

where the sum extends only over those frequencies contained within transition t. If we substitute equations of the form of (12-44) into (12-45), and perform the indicated summations, we obtain finally

$$\boldsymbol{\delta}\mathbf{Z}_t + \mathbf{A}_t\,\boldsymbol{\delta}\mathbf{n}_l + \mathbf{B}_t\,\boldsymbol{\delta}\mathbf{n}_u = \mathbf{C}_t \tag{12-46}$$

where $\mathbf{A}_t$ and $\mathbf{B}_t$ are full matrices.

Exercise 12-4: (a) Write explicit expressions for the elements of the matrices $\mathbf{T}_k$, $\mathbf{L}_k$, and $\mathbf{U}_k$, and the vector $\mathbf{R}_k$ in equation (12-42). (b) Assuming that the matrices $\mathscr{L}_k$, $\mathscr{U}_k$, and the vector $\mathscr{R}_k$ are *known*, write explicit expressions for the elements of $\mathbf{A}_t$, $\mathbf{B}_t$, and $\mathbf{C}_t$ in equation (12-46).

From equations (5-108) and (12-39) it follows that one can write

$$\boldsymbol{\delta}\mathbf{n}_m = \sum_t \mathbf{D}_{mt}\,\boldsymbol{\delta}\mathbf{Z}_t \tag{12-47}$$

where $\mathbf{D}_{mt}$ is a diagonal matrix with elements $(\mathbf{D}_{mt})_d = (\partial n_m/\partial Z_t)_d = (\mathscr{A}_d)^{-1}_{mj} - (\mathscr{A}_d)^{-1}_{mi}$, where $\mathscr{A}_d$ is the unperturbed rate matrix at depth-point d, and i and j are the lower and upper states in transition t. Using equation (12-47) in equation (12-46), we obtain the system

$$\begin{aligned}\mathbf{E}_t\,\boldsymbol{\delta}\mathbf{Z}_t &\equiv (\mathbf{I} + \mathbf{A}_t\,\mathbf{D}_{lt} + \mathbf{B}_t\,\mathbf{D}_{ut})\,\boldsymbol{\delta}\mathbf{Z}_t \\ &= \mathbf{C}_t - \mathbf{A}_t \sum_{t' \neq t} \mathbf{D}_{lt'}\,\boldsymbol{\delta}\mathbf{Z}_{t'} - \mathbf{B}_t \sum_{t' \neq t} \mathbf{D}_{ut'}\,\boldsymbol{\delta}\mathbf{Z}_{t'}\end{aligned} \tag{12-48}$$

with one such equation for each transition t.

The overall size of the system (12-48), which contains the full transition-to-transition coupling (interlocking) over all depth-points, is $(DT \times DT)$. A direct solution of these equations would thus require a time that scales as $T_D = c\,D^3T^3$, which for, say, $T \sim 20$ and $D \sim 50$ becomes impractically large. The system is therefore solved by iteration, using a *successive over-relaxation* (SOR) *method* (**526**, 438). In this approach there are *two* basic iteration cycles: (a) the SOR iteration, to obtain a definite set of $\boldsymbol{\delta}\mathbf{Z}_t$'s $(t = 1, \ldots, T)$ within a given stage of linearization, and (b) the overall linearization procedure, where successive sets of $\boldsymbol{\delta}\mathbf{Z}_t$'s are used to update rates, and the full statistical equilibrium equations are then re-solved. The SOR procedure is started by computing the solutions of the systems $\mathbf{E}_t\,\boldsymbol{\delta}\mathbf{Z}_t^{(0)} = \mathbf{C}_t$ $(t = 1, \ldots, T)$; this initial solution requires $c\,D^3T$ operations, and the resolved systems (equivalent to $\mathbf{E}_t^{-1}$) are saved. Then with any set of current estimates of the $\boldsymbol{\delta}\mathbf{Z}$'s, the righthand side of equation (12-48) can be evaluated for each transition in turn (note that only vector multiplications are involved, so the procedure is very fast); this yields a single vector of known value on the righthand side, and, using the previously resolved $\mathbf{E}_t$, a new value of $\boldsymbol{\delta}\mathbf{Z}_t$ is obtained. Each cycle in the SOR procedure requires $c\,D^2T^2$ operations, so if I iterations are necessary, the overall computing time scales as $T_{\text{SOR}} = c\,D^3T + c'I\,D^2T^2$, which is clearly favorable compared to T_D if $I < DT$.

By actual tests it is found (**37**) that this method works well, even though the SOR iterations resemble the equivalent-two-level-atom approach, in that only one transition at a time is treated. The reason is that this part of the calculation is required only to determine the $\boldsymbol{\delta}\mathbf{Z}$'s, which are merely one step of the overall linearization procedure (in itself designed to handle the interlocking problem self-consistently). Inasmuch as further steps in the linearization are presumed, the $\boldsymbol{\delta}\mathbf{Z}$'s at any given stage need not be known perfectly, but only with sufficient accuracy that the *error* in the current estimate of $\boldsymbol{\delta}\mathbf{Z}_t$ is smaller than the *full size* of the $\boldsymbol{\delta}\mathbf{Z}_t$'s of the next linearization step. In practice, the requirement $|\delta Z_t^{(i)} - \delta Z_t^{(i-1)}| < \varepsilon|\delta Z_t^{(i)}|$ (where i denotes the SOR iteration number) works well with ε set at about 10^{-2}. Both of the methods described in this section have proven to be very effective in a wide variety of physical problems, for different atoms, in stellar atmospheres of various types; we now turn to a discussion of some of the results obtained for early-type stars.

12-4 Light-Element Spectra in Early-Type Stars

From our study of the combined equations of transfer and statistical equilibrium, we have gained deep insight into the physics of spectral line-formation in stellar atmospheres. But it has also emerged that, when departures from LTE are taken into account, the equations to be solved become extremely intricate, and require special methods (which are computationally expensive).

It is therefore of considerable interest to answer the questions: "Do departures from LTE actually have a sensible effect in real stellar spectra?" "Are serious errors made in diagnostics of element abundances, or physical properties of the atmosphere, when non-LTE effects are ignored?" "Are there regions of the spectrum, portions of the H–R diagram, or classes of atoms and ions, for which we can conclude categorically that the use of an LTE approximation is safe, or is unsound?" From the practical point of view, it is important to know when it is adequate to assume LTE in the analysis of a stellar spectrum, for then the amount of computational work required is greatly reduced. Equally important is knowledge of when the results obtained from such analyses are *not* to be trusted. Answers to the questions posed above have come forth only quite recently as it became possible to obtain accurate numerical solutions of the non-LTE transfer equations, for elaborate multilevel atomic models in realistic model atmospheres, by applying on high-speed computers the techniques discussed in this chapter. It is now known that non-LTE effects are indeed of great consequence in several situations of basic interest and importance, and that a number of serious discrepancies between theory and observation are removed when departures from LTE are taken into account. A few of the results obtained thus far will be described below; further details can be found in the papers cited as well as reviews given in (**433**; **434**; **435**). It is not an exaggeration to state that work in this essential area has just begun, and much is yet to be learned from further efforts.

In many respects, the most fundamental spectrum in early-type stellar atmospheres is that of hydrogen. Not only do the hydrogen-line profiles serve as effective tools to measure basic stellar parameters, such as the effective temperature and surface gravity, but also, when one recognizes that hydrogen dominates the transfer of radiation through the major part of the spectrum in early-type stars, one realizes that a serious discrepancy between theory and observation for these lines has grave implications for the overall properties of the radiation field. Although very good agreement is usually obtained between predicted and observed H-line profiles for B- and A-type stars (cf. §10-5), strong disagreement has been found for the O-stars ($T_{\text{eff}} > 30{,}000°\text{K}$). The basic problem is that the H-lines are observed to have nearly-constant strength in the O-stars over the spectral range O9–O5 [as can be seen by inspection of the plates in (**465**) or (**5**)], while the LTE equivalent widths decrease markedly over the corresponding temperature range ($30{,}000°\text{K} \leqslant T_{\text{eff}} \leqslant 50{,}000°\text{K}$), owing to increased hydrogen ionization. This decrease can be offset somewhat by assuming a higher surface gravity in the models, but then inconsistencies arise. Typically it is found (**420**) that, if a match is made between the observed H-line *equivalent widths* and those calculated from LTE models, then the surface gravities required to obtain the fit are too large (by about a factor of 3) compared to those deduced from fundamental measurements of stellar masses and radii, or

from stellar-structure computations. Similarly, when detailed comparisons are made with line *profiles* (**510**), it is found that, if a fit is made to the line-wings, the line-core predicted by LTE is much too weak. These discrepancies vanish when departures from LTE are taken into account (**436**; **45**), and it is found that the non-LTE profiles are invariably much stronger than the LTE profiles, as shown in Figure 12-9 for a typical model. A comparison between the observed equivalent widths of Hβ and both LTE and non-LTE calculations is shown in Figure 12-10. It can easily be seen there that the non-LTE results are in much better agreement with observation, and that the LTE equivalent widths are systematically too small by factors of 3 to 5 in the extreme cases [see also (**174**)]. Further, an excellent fit can be obtained to observed profiles by using the non-LTE computations; this can be seen by comparing the fits displayed in (**510**) and (**45**), both of which use the same observational data. In sum, by taking into account departures from LTE, *major* improvements are obtained in matching observed H-line profiles for the O-stars. For the B-stars, the hydrogen lines are much less strongly affected by deviations from LTE, and for a line such as Hγ, LTE actually provides a very good approximation. Some important effects are found for Hα (**511**; **430**), where departures from LTE produce a deeper line-core, and a shallower wing than given by LTE. The latter effect arises because at the depths where the Hα-wing is formed, $b_3 > b_2$ (cf. §7-5) and therefore $S_{23} > B(\nu_{23}, T)$; the predicted changes have been confirmed by observation.

After hydrogen, helium is the next most important element in stellar atmospheres, and is represented by the He I spectrum in B-stars, and by both the He I and He II spectra in O-stars. These lines may be used to

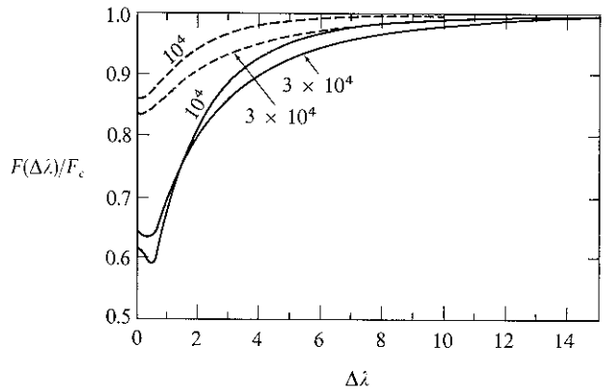

FIGURE 12-9
Hα profiles for O-star models with $T_{\text{eff}} = 45{,}000°\text{K}$; curves are labeled with surface gravity g. *Abscissa* gives displacement from line-center in Å; *ordinate* gives residual flux. *Solid curves*: non-LTE profiles; *dashed curves*: LTE profiles. Note marked strengthening of line by non-LTE effects. From (**45**), by permission.

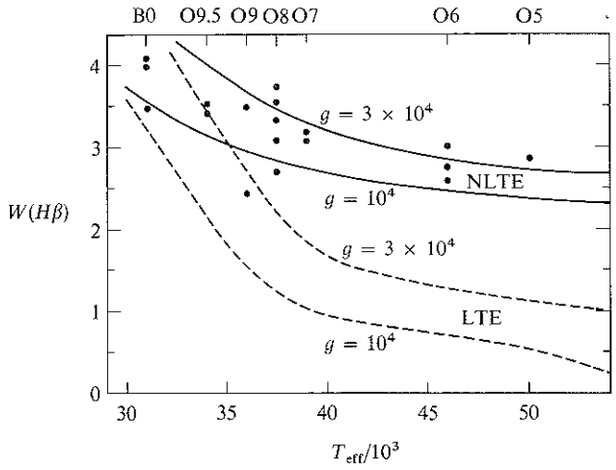

FIGURE 12-10
$H\beta$ equivalent widths in O-stars. *Solid curves*: non-LTE calculations; *dashed curves*: LTE calculations. *Dots*: observed values (**693**). *Ordinate* gives equivalent width in Å; *abscissae* give ($T_{eff}/10^3$) and spectral types.

estimate the helium abundance $N(\text{He})/N(\text{H})$, and in the O-stars, the ratio of strengths of the He II lines to He I lines provides a sensitive temperature-indicator. LTE calculations of He I lines in B-stars have produced excellent agreement with the observations for the blue-violet lines [see, e.g. (**390**; **480**; **481**)], especially when the most accurate line-broadening theories available are used (**438**; **439**). However, lines in the yellow-red regions of the spectrum [e.g., λ5876 ($2p\ ^3P$–$3d\ ^3D$) and λ6678 ($2p\ ^1P$–$3d\ ^1D$)] are always found to be stronger in B-star spectra than given by LTE computations, using the same model and abundance that fit the blue-violet lines. Formally these lines indicate abundances almost three times larger than that given by the blue-violet lines. When a detailed non-LTE calculation is made (**46**), using a realistic atomic model (see Figure 12-11), it is found that the lines mentioned above are strengthened dramatically, and that good agreement with observation is obtained, as shown in Figure 12-12 for λ6678. It is interesting that the largest non-LTE effects are found for the lines with the smallest values of $h\nu$, for, as mentioned in §5-3, these lines, with relatively small values of $h\nu/kT$, have collisional rates that are comparable to the radiative rates, and the classical argument is that these lines should therefore be in LTE. The reason for the large effects can be seen by examining the source function in the limit $\delta \equiv h\nu/kT < 1$. If we write $(b_l/b_u) = (1 + \beta)$, then

$$(S_l/B_\nu) = (e^{h\nu/kT} - 1)/[(b_l/b_u)e^{h\nu/kT} - 1]$$
$$= (e^{\delta} - 1)/[e^{\delta}(1 + \beta) - 1] \approx \delta/(\beta + \delta) = (1 + \beta/\delta)^{-1}$$

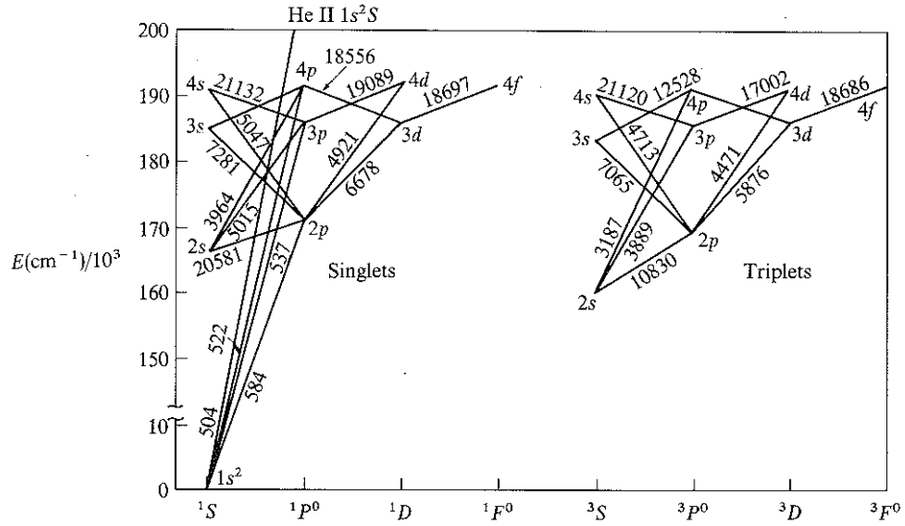

FIGURE 12-11
Model He I atom used in non-LTE calculation of He I spectrum. Ordinate gives wavenumber in cm^{-1} times 10^{-3}. All the transitions shown were included in the calculation explicitly as "primary" transitions in the sense defined in §12-3; all other dipole-permitted lines up through $n = 10$ were treated as "secondary" transitions, with rates fixed by an equivalent-two-level-atom computation. The singlets and triplets are treated separately. From (**46**), by permission.

which shows that a *given* fractional departure β of the level-population ratio is *amplified* by a factor δ^{-1} when $\delta < 1$. The occupation numbers, and hence the ratio (b_l/b_u) will, in general, be determined by processes *other than* those in the line. In the O- and B-stars the level populations are set by photoionizations and recombinations, and thus β will generally be nonzero despite a large collisional rate in a line. Once $\beta \neq 0$, large departures of S_l from B_ν become possible; the same argument, by the way, explains why even tiny departures from LTE have very large effects in the radio-frequency recombination lines observed in nebulae (**199**). In the O-stars, LTE predicts too-weak lines for both He I and He II. If one attempts to match the observed equivalent widths, without regard to line-profiles (**420**), the derived helium abundances are about a factor of two higher than the value obtained from B-stars, or from the nebulae in the interstellar medium from which the O-stars have just formed. If a fit is made to the wings of the line profiles (**510**), it is again found that the line cores are much too weak. On the other hand, a non-LTE calculation (**45**) yields excellent agreement with observed line strengths and profiles at the "standard" abundance $N(\text{He})/N(\text{H}) = 0.1$ [see also (**174**; **175**)].

Departures from LTE on occasion introduce *large* errors into abundance estimates based on LTE calculations. For example, LTE determinations of

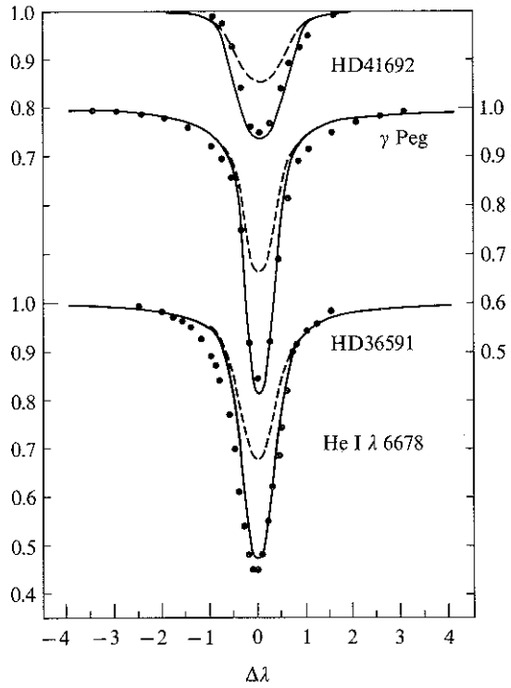

FIGURE 12-12
Comparison of observed (**390**) He I λ6678 profiles (*dots*) with non-LTE (*solid curves*) and LTE (*dashed curves*) calculations for three B-stars. *Ordinate*: residual flux; *abscissa*: displacement from line-center in Å. From (**46**), by permission.

the abundance of Mg in O-stars are in error by a factor of 10 or more (**431**), and only when a non-LTE calculation is made, is a fit obtained to the observations, with an abundance near the solar (and accepted "cosmic") value, as shown in Figure 12-13 [see also (**384**; **586**)]. Similarly, LTE analyses of the spectrum of Ne I in B-stars have routinely given a neon abundance of 5×10^{-4} relative to hydrogen, in disagreement with the value 10^{-4} obtained from nebulae, the solar corona, the solar wind, and cosmic rays (both galactic and solar-produced). Again the discrepancy is removed when a non-LTE calculation is done (**47**) and it is found that the Ne I spectrum, consisting primarily of lines in the red, is affected by essentially the same mechanism as that described above for He I. Finally, the spectra of Si III and Si IV in B- and O-stars show significant non-LTE effects (**349**). Departures from LTE increase the computed line-strengths by 50 to 70 percent, and line-core intensities decrease by a factor of 0.6 relative to LTE. To

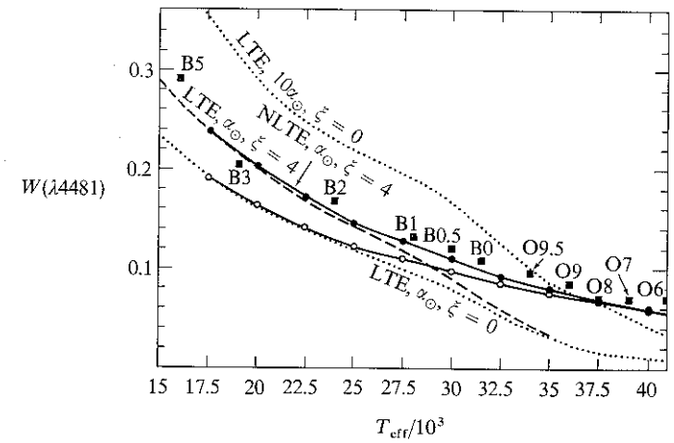

FIGURE 12-13
Strength of Mg II λ4481 in B- and O-stars. *Ordinate*: equivalent width in Å; *abscissa*: $(T_{\text{eff}}/10^3)$. *Squares*: observed values at the indicated spectral types. *Dotted curves*: LTE calculations at a solar or ten-times solar Mg abundance, with zero microturbulence. *Dashed curve*: solar Mg abundance, and 4 km s^{-1} microturbulence. *Solid curves*: non-LTE calculation at solar Mg abundance, and zero (open circles) or 4 km s^{-1} (filled circles) microturbulence. Note that for the O-stars the choice of microturbulent velocity does not influence the result, and that the LTE predictions yield much too small a line-strength, while the non-LTE results agree well with the observations.

produce lines of similar strength in LTE, an abundance 5 to 7 times the solar value would be required. In short, there is ample evidence that, at least for early-type stars, deviations from LTE may produce large and important changes in abundance estimates.

Yet another context in which non-LTE effects enter preeminently in early-type stellar spectra is in the production of emission lines. For example, in the Of stars, the lines of He II λ4686, C III λ5696, and N III λλ4634–41 are seen in emission, while other lines of the same ionic spectra are in absorption. In the case of He II λ4686, studies have been made of selective excitation mechanisms (**45**) and the effects of atmospheric extension in static models (**376**); while weak emission lines have been obtained in such work, it is clear that a quantitative fit to the observed line-strengths will be obtained only when large-scale atmospheric expansion is taken into account (see Chapters 14 and 15). The C III line has not yet been subjected to analysis. For the N III lines there exists a subgroup of the Of stars, designated O((f)), in which the N III lines are weakly in emission and the He II line is in absorption (**657**); these appear to be near-main-sequence objects in which the effects of atmospheric extension and expansion, if present at all, may be ignored.

A detailed non-LTE study of the N III doublet spectrum (**429**; **440**; **443**; **433**) gives a plausible explanation of how the $\lambda\lambda$4634–41 lines from the $3p\ ^2P \rightarrow 3d\ ^2D$ transition can be *in emission*, while the next lower multiplet, $\lambda\lambda$4097, 4103 ($3s\ ^2S \rightarrow 3p\ ^2P$) is *in absorption*. The two essential ingredients required to produce this result are (a) a mechanism for populating the $3d$ state with enough electrons to induce the $3d \rightarrow 3p$ emission, and (b) a way of depopulating the $3p$ state so that emission does not appear in the $3p \rightarrow 3s$ transition. Both the $3s$ and $3d$ states can be populated directly by transitions from the ground state (see Figure 12-14) and, when the ultraviolet radiation field at λ374 and λ452 is intense, these states will have large occupation numbers. In *static* atmospheres, however, the resonance lines come into detailed balance, and therefore do not lead to a very large overpopulation of the $3d$ state. Rather, the overpopulation results from *dielectronic recombination* from the $2s2p(^1P)3d$ state—which, for the N^{++} ion, just happens to lie barely above the ionization limit. Stabilization of this state occurs when the $2p$ electron drops to the $2s$ level, and proceeds at a very rapid rate, feeding electrons directly into the $3d$ state. These electrons then decay $3d \rightarrow 3p$, producing the emission. Further, it turns out that, because of the particular structure of the N^{++} ion, the most probable route of exit from $3p$ is *not* $3p \rightarrow 3s$, but rather via "two-electron jumps" of the form $2s^23p \rightarrow 2s2p^2$ (see Figure 12-14); the latter process occurs so efficiently that it drains the $3p$ state, leaving the $3s \rightarrow 3p$ line in absorption. Inasmuch as the $\lambda\lambda$4634–

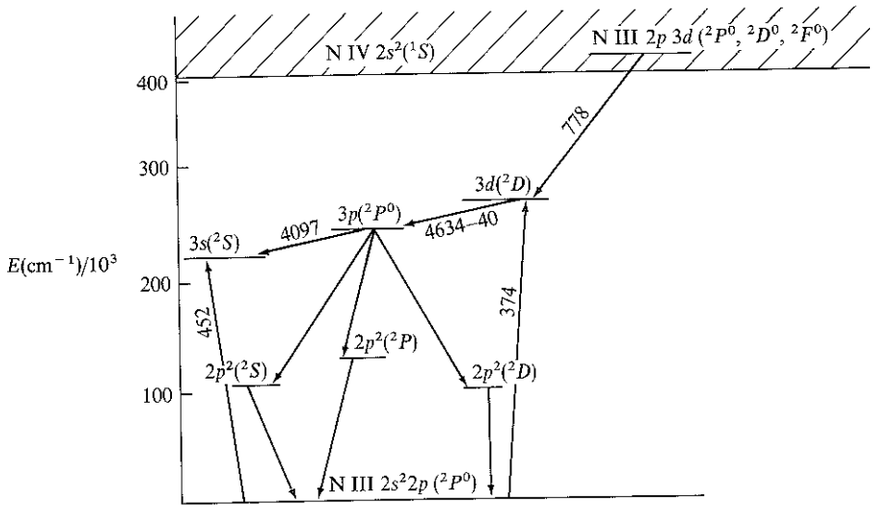

FIGURE 12-14
Simplified term diagram for the lower states of N^{++}, as well as the autoionizing state $2s2p(^1P)3d$ from which dielectronic recombinations can occur directly into the $2s^23d$ state, leading to emission at $\lambda\lambda$4634–40.

41 emission lines can be produced in planar, static atmospheres, solely as a result of the atomic structure of the ion responsible, they must be regarded as *intrinsic* emission lines. In *expanding* atmospheres, direct pumping of the $3s$ and $3d$ states will occur in the $2p \rightarrow 3s$ and $2p \rightarrow 3d$ transitions, because the resonance lines will be Doppler shifted into the bright adjacent continuum and will not be in detailed balance. This greatly enhances the $3d \rightarrow 3p$ emission (thus explaining the very bright emission seen in the Of stars, which are known to have expanding envelopes and stellar winds), while the increased $3s$ population, coupled with the drain from $3p$, assures that the $3s \rightarrow 3p$ lines remain in absorption. The calculations show (see Figure 12-15) that the $\lambda\lambda$4634–40 lines make the transition from absorption to emission

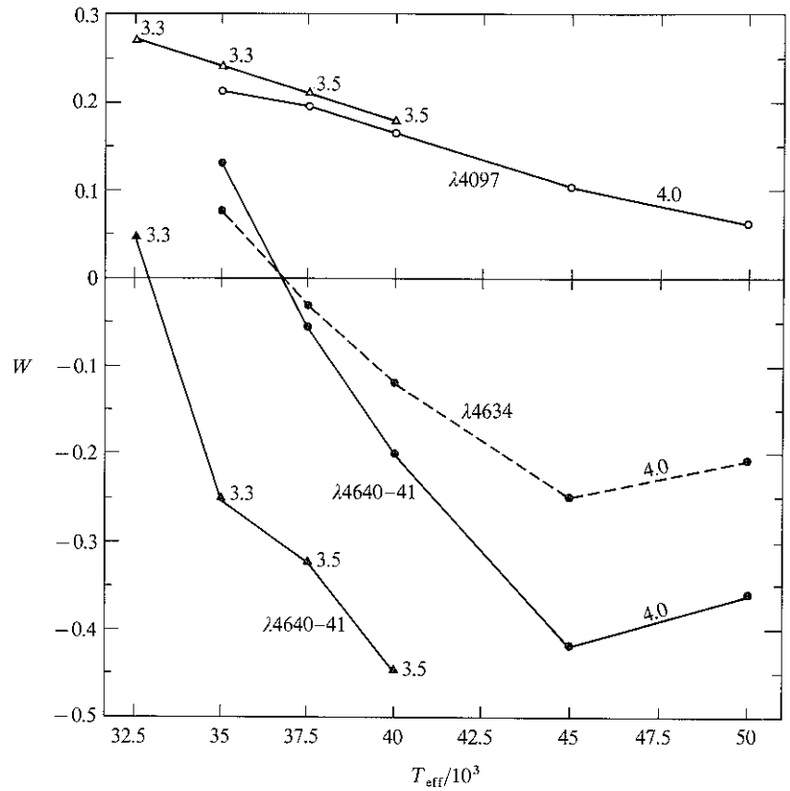

FIGURE 12-15
N III line-strengths for the $3s \rightarrow 3p$ (λ4097) and $3p \rightarrow 3d$ ($\lambda\lambda$4634–40) lines in planar, static, model atmospheres. *Ordinate*: equivalent width (negative values denote emission); *abscissa*: ($T_{\text{eff}}/10^3$). Note that the λ4097 line remains in absorption while $\lambda\lambda$4634–40 make a transition from absorption to emission at a spectral type near O6. From (**443**) by permission.

at about spectral type O6 near the main sequence, in agreement with observation (**443**).

Many more analyses, for a wide variety of atoms and stellar atmospheres, need yet to be done (particularly for later spectral types) before the questions posed at the beginning of this section can be answered fully, and efforts in this direction will be richly rewarded with interesting results.

13

Line Formation with Partial Frequency Redistribution

In the process of *scattering* in spectral line-formation, an atom is excited from one bound level to another by the absorption of a photon, and then decays radiatively back to the original state, with the emission of a photon. In our work on line-transfer we have, thus far, assumed that scattering is either strictly coherent, or that the photons are completely redistributed over the line profile. Neither of these limits is achieved exactly in stellar atmospheres, and it is necessary to consider the redistribution of photons, in angle and frequency, in some detail, and to calculate *redistribution functions*, which describe the scattering process precisely. This calculation proceeds in two steps. We first consider a single atom in its own frame of reference, and compute the form of any redistribution that occurs within the substructure of the bound states. Then, recognizing that what is actually observed in a stellar atmosphere is an entire ensemble of atoms moving with a thermal velocity distribution, we take into account the Doppler redistribution in frequency produced by the atoms' motion. Doppler redistribution arises because the incident and emergent photons travel, in general, in different directions; in this event the projection of the atom's velocity vector along the propagation vectors will be different for the two photons, and a differential

Doppler shift must occur. The final redistribution function is obtained by averaging over all possible velocities. Once the redistribution function is known, a transfer calculation can be carried out, taking into full account any correlations (or lack thereof) between incoming and outgoing photon frequencies.

In this chapter the notation and methods of Hummer (**313**) will be adopted; the reader is referred to this paper for further details and other results. An interesting analysis of the problem from a rather different conceptual point of view can be found in (**295**) and (**296**).

13-1 Redistribution in the Atom's Frame

Let us first consider the nature of redistribution in the *rest frame of the atom.* Let the frequency displacement from line center of the incoming photon, *measured in the atom's frame*, be denoted by ξ', its direction by $\mathbf{n}'$, and the frequency displacement and direction of the outgoing photon by ξ and $\mathbf{n}$. Assume that the material has no preferred directions on an atomic scale so that the atomic absorption profile $f(\xi')$ is isotropic; f is normalized such that $\int_{-\infty}^{\infty} f(\xi')\, d\xi' = 1$. Further, suppose that the *frequency redistribution function* $p(\xi', \xi)$ gives the probability that a photon absorbed in the frequency range $(\xi', \xi' + d\xi')$ is emitted into the range $(\xi, \xi + d\xi)$, while the *angular phase function* $g(\mathbf{n}', \mathbf{n})$ describes the probability that the photon is scattered from solid angle $d\omega'$ in direction $\mathbf{n}'$, into solid angle $d\omega$ in direction $\mathbf{n}$. These functions are normalized such that

$$\int_{-\infty}^{\infty} p(\xi', \xi)\, d\xi' = \int_{-\infty}^{\infty} p(\xi', \xi)\, d\xi = 1 \tag{13-1}$$

and

$$(4\pi)^{-1} \oint g(\mathbf{n}', \mathbf{n})\, d\omega' = (4\pi)^{-1} \oint g(\mathbf{n}', \mathbf{n})\, d\omega = 1 \tag{13-2}$$

The phase functions most useful in describing atomic scattering are those for isotropic and dipole scattering, given by equations (2-17) and (2-18).

In terms of the functions just defined, we may now write the probability that a photon $(\xi', \mathbf{n}')$ is absorbed as $f(\xi')\, d\xi'\, d\omega'/4\pi$, and the probability that, *if* a photon $(\xi', \mathbf{n}')$ is absorbed, *then* a photon $(\xi, \mathbf{n})$ is emitted as $p(\xi', \xi)g(\mathbf{n}', \mathbf{n})\, d\xi\, d\omega/4\pi$. Thus the *joint probability* that a photon $(\xi', \mathbf{n}')$ is absorbed *and* a photon $(\xi, \mathbf{n})$ is emitted is $f(\xi')p(\xi', \xi)\, d\xi'\, d\xi\, g(\mathbf{n}', \mathbf{n})(d\omega'/4\pi)(d\omega/4\pi)$.

We must now specify the functions $f(\xi')$ and $p(\xi', \xi)$. Following Hummer, we shall consider the following four categories: (a) Case I, zero line width; (b) Case II, radiation damping in the upper state, and coherence in the atom's rest frame; (c) Case III, complete redistribution in atom's frame; (d) Case IV, subordinate-line redistribution between two broadened states. Let us examine these in turn.

(a) Case I. Here we consider an idealized atom with two perfectly sharp states. Then $f(\xi')\, d\xi' = \delta(\xi' - \xi_0)\, d\xi'$, and $p(\xi', \xi) = \delta(\xi' - \xi)$ where δ denotes the Dirac function, and ξ_0 is the line-center frequency; clearly there is no redistribution in the atom's frame in this case. It is obvious that the conditions described here do not apply to any real line, for normally one (or both) levels will be broadened. It is useful, nevertheless, to study this limiting case, for it demonstrates the effects of Doppler redistribution *alone*, as seen by an observer in the laboratory frame examining the ensemble of moving atoms.

(b) Case II. Here we envision an atom with a perfectly sharp lower state, and an upper state whose finite lifetime against radiative decay (back to the lower state) leads to a Lorentz profile

$$f(\xi') = (\delta/\pi)/[(\xi' - \xi_0)^2 + \delta^2] \tag{13-3}$$

where $\delta \equiv \Gamma_R/4\pi$, and Γ_R is the radiative damping width of the upper state. We assume that there are no additional perturbations of the atom while it is in its upper state; then there will be no reshuffling of electrons among substates of the upper state, and the decay to the lower state will produce a photon of exactly the same frequency as was absorbed. Thus we again have $p(\xi', \xi) = \delta(\xi - \xi')$. This case applies to resonance lines in media of such low densities that collisonal broadening of the upper state is completely negligible—for example, the Lyman α line of hydrogen in the interstellar medium.

(c) Case III. The basic physical picture here is of an atom with a perfectly sharp lower state, and a broadened upper state, in a medium where collisions are so frequent that *all* excited electrons are randomly reshuffled over the substates of the upper state before emission occurs. The absorption profile is again the Lorentz profile given by equation (13-3), where δ now represents the full width (radiative plus collisional) of the upper state. In this extreme limit, the frequency of the emitted photon will have *no* correlation with the frequency of that absorbed; the probability for emission at any particular frequency is then proportional to the number of substates present at that frequency, and hence to the absorption profile itself. When complete redistribution in the atom's frame occurs we thus have

$$p(\xi', \xi)\, d\xi = f(\xi)\, d\xi = (\delta/\pi)\, d\xi/[(\xi - \xi_0)^2 + \delta^2] \tag{13-4}$$

which shows clearly that $p(\xi', \xi)$ is independent of ξ', and that the joint probability of absorption at ξ' and emission at ξ is proportional to $f(\xi')f(\xi)$.

(d) Case IV. Here we suppose that the line is formed by an absorption from a broadened state i, to a broadened upper state j, followed by a radiative decay to state i. This picture is appropriate to scattering in subordinate lines. Because the electron returns to the same level as that from which it was

excited, the entire circuit is treated as a single quantum-mechanical process, and an expression is obtained for the *product* $f(\xi')p(\xi', \xi)$. An expression for the joint probability of absorption and emission was derived by Weisskopf (**663**) and by Woolley (**683**). A subsequent analysis by Heitler (**293**, 198) yielded a different result that was widely quoted in the astrophysical literature, but is now known to be in error (**489**, 195); the earlier formula is, in fact, correct. The derivation is straightforward but lengthy, so only the final expression will be quoted here, namely

$$\begin{aligned} &f(\xi')p(\xi', \xi) \\ &= (\delta_i^2\,\delta_j/\pi^2)4(\delta_i + \Delta)/\{[(\xi' - \xi_0)^2 + \Delta^2][(\xi - \xi_0)^2 + \Delta^2][(\xi' - \xi)^2 + 4\delta_i^2]\} \\ &\quad + (\delta_i\,\delta_j/\pi^2)/\{[(\xi - \xi_0)^2 + \Delta^2][(\xi - \xi')^2 + 4\delta_i^2]\} \\ &\quad + (\delta_i\,\delta_j/\pi^2)/\{[(\xi - \xi')^2 + 4\delta_i^2][(\xi' - \xi_0)^2 + \Delta^2]\} \\ &\quad + (\delta_i^2/\pi^2)/\{[(\xi' - \xi_0)^2 + \Delta^2][(\xi - \xi_0)^2 + \Delta^2]\} \end{aligned} \tag{13-5}$$

where $\Delta \equiv \delta_i + \delta_j$ [see (**684**, 164–168) for a detailed derivation and discussion].

It is easy to show that $f(\xi')p(\xi', \xi) = f(\xi)p(\xi, \xi')$, as would be expected in the symmetrical process $i \rightarrow j \rightarrow i$. Further, by inspection of the denominators in equation (13-5), it is easy to see that for a given value of ξ, there are two relative maxima, one with $\xi' = \xi$ and the other with $\xi' = \xi_0$. These can be understood physically as follows. (1) Most absorptions will arise from the center of the lower level; when these excite a particular substate of the upper state, the decay occurs most often back to the center of the lower level. It is clear that ξ will then equal ξ', and that this process occurs with a relatively high probability. (2) Alternatively, there is a very high probability that an atom will be excited from the center of the lower state to the center of the upper by photons with $\xi' = \xi_0$; the upper state may then decay to any arbitrary substate of the lower, in particular, yielding the frequency ξ. The inverse also occurs: a transition from an arbitrary substate of the lower level has high probability of exciting the center substate of the upper level, and the most probable decay is back to the center of the lower level. Thus there will be a peak in emission at line center (i.e., at $\xi = \xi_0$), and indeed one sees that there is such a peak in equation (13-5).

The laboratory-frame redistribution function corresponding to equation (13-5) has not appeared in the literature [owing to use of the erroneous result of Heitler], though it could be derived straightaway (the calculation would be tedious, however). Moreover, the formulation of a correct transfer equation for subordinate-line redistribution is quite complicated; we shall not, therefore, consider Case IV further in what follows.

Of the four cases defined above, Cases II and III are of the greatest astrophysical importance. In fact, neither of these extreme limits is attained, and

the typical situation of interest for resonance-line formation is one in which the upper state is broadened by *both* radiation damping and elastic collisions, with damping-widths δ_R and δ_C respectively. In this case, one would expect the line profile to be given by equation (13-3) with $\delta = \delta_R + \delta_C$. Of the atoms in the upper state, a fraction $\gamma \equiv \delta_R/(\delta_R + \delta_C)$ would be expected to decay radiatively, and hence to emit coherently in the atom's frame (recall the lower state is perfectly sharp). The remaining fraction $(1 - \gamma) = \delta_C/(\delta_R + \delta_C)$ would have suffered collisions, and would be expected to have been completely redistributed. Thus we could now write

$$p(\xi', \xi) = \gamma\,\delta(\xi' - \xi) + (1 - \gamma)(\delta/\pi)/[(\xi - \xi_0)^2 + \delta^2] \tag{13-6}$$

This result was derived by Zanstra (**690**; **691**), who treated the radiating atom as a classical oscillator. A detailed quantum-mechanical calculation (**489**) recovers equation (13-6) when the lower state is presumed to be perfectly sharp. If *inelastic* collisions occur with sufficient frequency, δ_I, to contribute to the total width of the state, then $\delta = \delta_R + \delta_C + \delta_I$; we now write $\gamma \equiv (\delta_R + \delta_I)/(\delta_R + \delta_C + \delta_I)$, for it is only the elastic collisions that reshuffle atoms among upper-level substates (**489**). In this case it is also necessary to introduce an additional emission source accounting for collisional excitations of the upper level (see §13-4). In sum, we see that the redistribution function in this more general situation can be expressed as a linear combination of the results for Cases II and III.

13-2 Doppler-Shift Redistribution in the Laboratory Frame

GENERAL FORMULAE

Let us now consider the effects of the Doppler shifts introduced by the motion of the scattering atoms relative to the laboratory frame. In this section we shall derive expressions that describe the full *angular and frequency dependence* of redistribution in the scattering process. In practice, relatively few radiative transfer calculations of relevance to line formation in stellar atmospheres have been done with the scattering treated in this much detail (the dimensionality of the problem is large!), and the *angle-averaged* functions to be derived in §13-3 are generally much more useful in application. Following Hummer's treatment, we first deduce general formulae for the observer's-frame redistribution function, and then calculate explicit results for the specific cases defined above. As was discussed in §2-1, the redistribution function $R(\nu', \mathbf{n}'; \nu, \mathbf{n})\,d\nu'\,d\nu(d\omega'/4\pi)(d\omega/4\pi)$ gives the joint probability of scattering a photon from a laboratory frame frequency $(\nu', \nu' + d\nu')$, and solid angle $d\omega'$ in direction $\mathbf{n}'$, into frequency $(\nu, \nu + d\nu)$, and solid angle $d\omega$

in direction $\mathbf{n}$. The function is normalized such that [equation (2-7)]

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

Suppose an atom moving with velocity $\mathbf{v}$, which remains fixed during the scattering process, absorbs a photon $(v', \mathbf{n}')$ and emits a photon $(v, \mathbf{n})$, as measured in the laboratory frame. Neglecting the aberration of directions in transforming from the atom's frame to the laboratory frame, the corresponding atom's frame frequencies for the absorption and emission are

$$\xi' = v' - v_0(\mathbf{v} \cdot \mathbf{n}')/c \tag{13-7a}$$

and

$$\xi = v - v_0(\mathbf{v} \cdot \mathbf{n})/c \tag{13-7b}$$

As was shown in §13-1, the joint probability of absorption of a photon $(\xi', \mathbf{n}')$ with subsequent emission of a photon $(\xi, \mathbf{n})$ measured in the atom's frame is $f(\xi')p(\xi', \xi)g(\mathbf{n}', \mathbf{n})\, d\xi'\, d\xi(d\omega'/4\pi)(d\omega/4\pi)$. Transforming this expression to the laboratory frame via equations (13-7) we can write

$$R_v(v', \mathbf{n}'; v, \mathbf{n}) = f(v' - v_0\mathbf{v} \cdot \mathbf{n}'/c)p(v' - v_0\mathbf{v} \cdot \mathbf{n}'/c, v - v_0\mathbf{v} \cdot \mathbf{n}/c)g(\mathbf{n}', \mathbf{n}) \tag{13-8}$$

where the subscript v implies that the redistribution is produced by an atom of velocity $\mathbf{v}$. To find the net result for the entire ensemble of atoms, we must average over the velocity distribution, which is assumed to be Maxwellian. To perform this average, introduce an orthogonal triad of reference axes $(\mathbf{n}_1, \mathbf{n}_2, \mathbf{n}_3)$, with $\mathbf{n}_1$ and $\mathbf{n}_2$ chosen to be coplanar with $\mathbf{n}'$ and $\mathbf{n}$, and with $\mathbf{n}_1$ bisecting the angle Θ between them (see Figure 13-1). Then we may write

$$\mathbf{n}' = (\cos \tfrac{1}{2}\Theta)\mathbf{n}_1 + (\sin \tfrac{1}{2}\Theta)\mathbf{n}_2 = \alpha\mathbf{n}_1 + \beta\mathbf{n}_2 \tag{13-9}$$

and

$$\mathbf{n} = (\cos \tfrac{1}{2}\Theta)\mathbf{n}_1 - (\sin \tfrac{1}{2}\Theta)\mathbf{n}_2 = \alpha\mathbf{n}_1 - \beta\mathbf{n}_2 \tag{13-10}$$

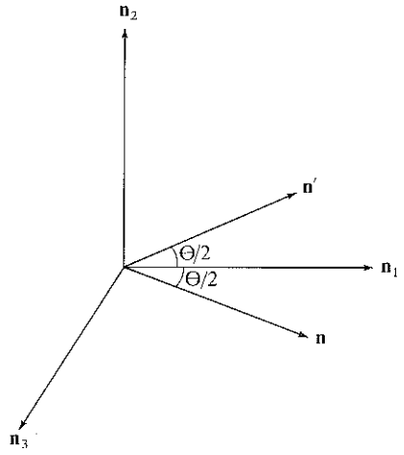

FIGURE 13-1
Coordinate axes used in calculation of redistribution function. The vectors $\mathbf{n}_1$, $\mathbf{n}_2$, $\mathbf{n}$, and $\mathbf{n}'$ are coplanar. The vector $\mathbf{n}_1$ bisects the angle θ, $(0 \leq \theta \leq \pi)$, between $\mathbf{n}'$ and $\mathbf{n}$.

Resolve v along these axes, and write $v_i = \mathbf{v} \cdot \mathbf{n}_i$; for convenience express velocities in dimensionless thermal units

$$\mathbf{u} = \mathbf{v}/v_{\text{thermal}} = (m_A/2kT)^{\frac{1}{2}}\mathbf{v} \tag{13-11}$$

where m_A is the mass of an atom, and introduce the Doppler width

$$w = (v_0/c)(2kT/m_A)^{\frac{1}{2}} = v_0(v_{\text{thermal}}/c) \tag{13-12}$$

Then the Maxwellian velocity distribution [equation (5-1)] is

$$P(u_1, u_2, u_3)\, du_1\, du_2\, du_3 = \pi^{-\frac{3}{2}} \exp[-(u_1^2 + u_2^2 + u_3^2)]\, du_1\, du_2\, du_3 \tag{13-13}$$

We now average equation (13-8) against the velocity distribution of equation (13-13) to obtain

$$\begin{aligned} R(v', \mathbf{n}'; v, \mathbf{n}) &= \int_{-\infty}^{\infty} du_1 \int_{-\infty}^{\infty} du_2 \int_{-\infty}^{\infty} du_3\, P(u_1, u_2, u_3) R_u(v', \mathbf{n}'; v, \mathbf{n}) \\ &= \pi^{-1} g(\mathbf{n}', \mathbf{n}) \int_{-\infty}^{\infty} du_1\, e^{-u_1^2} \int_{-\infty}^{\infty} du_2 \\ &\quad \times e^{-u_2^2} f[v' - w(\alpha u_1 + \beta u_2)] \\ &\quad \times p[v' - w(\alpha u_1 + \beta u_2), v - w(\alpha u_1 - \beta u_2)] \end{aligned} \tag{13-14}$$

where the integration over u_3 has been carried out explicitly. An alternative form for R, which will prove useful, can be derived by choosing $\mathbf{n}_1$ to lie along $\mathbf{n}'$. Then $\mathbf{v} \cdot \mathbf{n}' = v_1$, while

$$\mathbf{v} \cdot \mathbf{n} = v_1 \cos \Theta + v_2 \sin \Theta = \alpha' v_1 + \beta' v_2 \tag{13-15}$$

We then find

$$\begin{aligned} R(v', \mathbf{n}'; v, \mathbf{n}) &= \pi^{-1} g(\mathbf{n}', \mathbf{n}) \int_{-\infty}^{\infty} du_1\, e^{-u_1^2} f(v' - wu_1) \\ &\quad \times \int_{-\infty}^{\infty} du_2\, e^{-u_2^2} p[v' - wu_1, v - w(\alpha' u_1 + \beta' u_2)] \end{aligned} \tag{13-16}$$

With the aid of these general formulae, we can now compute the redistribution functions for the various cases defined in §13-1. Note first, however, that in both Cases I and II the scattering is coherent in the atom's frame, so that

$$p(\xi', \xi) = \delta(\xi' - \xi) = \delta(v - v' + 2w\beta u_2) \tag{13-17}$$

In these cases u_1 no longer appears explicitly in the expression for p, and the integrations written above are simplified. Substituting equation (13-17) into (13-14), the integration over u_2 can now be performed explicitly (transforming to the variable $z \equiv 2w\beta u_2$ in order to preserve normalization of the Dirac

δ-function). We then obtain, for $\beta \neq 0$,

$$R_c(\nu', \mathbf{n}'; \nu, \mathbf{n}) = \frac{g(\mathbf{n}', \mathbf{n})}{2\pi\beta w} \exp\left[-\frac{(\nu - \nu')^2}{4\beta^2 w^2}\right] \times \int_{-\infty}^{\infty} e^{-u^2} f\left[\tfrac{1}{2}(\nu + \nu') - w\alpha u\right] du \qquad (13\text{-}18)$$

while if $\beta = 0$ we find

$$R_c(\nu', \mathbf{n}'; \nu, \mathbf{n}) = \pi^{-\frac{1}{2}} g(\mathbf{n}', \mathbf{n})\, \delta(\nu' - \nu) \int_{-\infty}^{\infty} e^{-u^2} f(\nu' - wu)\, du \qquad (13\text{-}19)$$

Here the subscript c denotes coherence in the atom's frame.

Exercise 13-1: Verify equations (13-18) and (13-19).

Finally, we note that the results can be written in a concise and convenient form by expressing frequency displacements from line center in Doppler units:

$$x \equiv (\nu - \nu_0)/w \qquad (13\text{-}20)$$

and

$$x' \equiv (\nu' - \nu_0)/w \qquad (13\text{-}21)$$

and writing

$$R(x', \mathbf{n}'; x, \mathbf{n}) \equiv R(\nu', \mathbf{n}'; \nu, \mathbf{n})(d\nu'/dx')(d\nu/dx) = w^2 R(\nu', \mathbf{n}'; \nu, \mathbf{n}) \qquad (13\text{-}22)$$

to yield normalization of $R(x', \mathbf{n}'; x, \mathbf{n})$ when integrated over x' and x.

RESULTS FOR SPECIFIC CASES

(a) Case I. If we substitute $f(\xi') = \delta(\xi' - \nu_0)$ into equation (13-18), we obtain immediately

$$R_I(\nu', \mathbf{n}'; \nu, \mathbf{n}) = \frac{g(\mathbf{n}', \mathbf{n})}{2\pi\alpha\beta w^2} \exp\left[\frac{-(\nu - \nu')^2}{4\beta^2 w^2}\right] \exp\left[\frac{-(\nu + \nu' - 2\nu_0)^2}{4\alpha^2 w^2}\right] \qquad (13\text{-}23)$$

Now noting that $2\alpha\beta = 2\sin(\frac{1}{2}\Theta)\cos(\frac{1}{2}\Theta) = \sin\Theta$, that $\alpha^2 + \beta^2 = 1$, and transforming to dimensionless units via equations (13-20) through (13-22), we have

$$R_I(x', \mathbf{n}'; x, \mathbf{n}) = [g(\mathbf{n}', \mathbf{n})/\pi \sin\Theta] \exp[-x^2 - (x' - x\cos\Theta)^2 \csc^2\Theta] \qquad (13\text{-}24)$$

Exercise 13-2: Fill in the steps required to obtain equation (13-24).

From this result, we immediately notice that, *even though the scattering in the atom's frame was presumed to be strictly coherent, there is significant redistribution in the laboratory frame.* This is a point of very great physical significance, as was stressed by Thomas (**622**). He examined the frequency dependence of scattered radiation [cf. equation (2-9)] using several simple analytical expressions for the incident intensity, and showed that in the Doppler core ($x \lesssim 3$), over which the absorption profile varies by a factor of 10^4, the emission profile departs from the absorption profile by less than a factor of four! Thus, *within the Doppler core, the assumption of complete redistribution of the scattered radiation provides an excellent approximation* (it will be easy to see this using the angle-averaged redistribution functions derived in §13-2; see also Figure 13-2). This result provided much of the motivation for assuming complete noncoherence in the treatment of line formation, as was done in Chapters 11 and 12. Outside the line core, complete redistribution is not achieved as closely, and one must use the correct redistribution function. For example, in the case considered here, a simple physical argument (**597**) shows that the radiation in the far line-wings is about two-thirds noncoherent, and one-third coherent. On the other hand, outside the Doppler core the line-wings will often be swamped by the continuum (unless the profile has a large damping parameter), in which event the details of redistribution become less significant.

(b) Case II. Here we substitute equation (13-3) into (13-18) to obtain

$$R_{II}(\nu', \mathbf{n}'; \nu, \mathbf{n}) = \frac{g(\mathbf{n}', \mathbf{n})}{2\pi\alpha\beta w^2} \exp\left[\frac{-(\nu - \nu')^2}{4\beta^2 w^2}\right] \left(\frac{\delta}{\pi\alpha w}\right) \times \int_{-\infty}^{\infty} e^{-u^2} \left[\left(\frac{\nu' + \nu - 2\nu_0}{2\alpha w} - u\right)^2 + \left(\frac{\delta}{\alpha w}\right)^2\right]^{-1} du \qquad (13\text{-}25)$$

Transforming to dimensionless frequency units via equations (13-20) through (13-22), and recalling the definition of the Voigt function $H(a, v)$ [see equation (9-34)], we obtain, finally

$$R_{II}(x', \mathbf{n}'; x, \mathbf{n}) = \frac{g(\mathbf{n}', \mathbf{n})}{\pi \sin\Theta} \exp[-\tfrac{1}{2}(x - x')^2 \csc^2 \tfrac{1}{2}\Theta] \times H[a \sec\tfrac{1}{2}\Theta, \tfrac{1}{2}(x + x')\sec\tfrac{1}{2}\Theta] \qquad (13\text{-}26)$$

where $a \equiv (\delta/w)$. Although this result is relatively complicated, efficient methods to evaluate $H(a, v)$ exist, and R_{II} can be calculated fairly easily. As will emerge in §13-3, R_{II} simulates nearly complete redistribution in the line core ($x \lesssim 2.5$), but becomes nearly coherent in the line wings.

Exercise 13-3: Fill in the missing steps leading to equation (13-26).

(c) Case III. Here, as was noted in equation (13-4), $p(\xi', \xi)$ is independent of ξ'. In this case equation (13-16) is particularly useful, and substituting from equations (13-3) and (13-4) we find

$$R_{\text{III}}(\nu', \mathbf{n}'; \nu, \mathbf{n}) = \frac{g(\mathbf{n}', \mathbf{n})}{\pi} \int_{-\infty}^{\infty} du_1\, e^{-u_1^2} \left(\frac{\delta}{\pi}\right) [(\nu' - \nu_0 - wu_1)^2 + \delta^2]^{-1}$$

$$\times \int_{-\infty}^{\infty} du_2\, e^{-u_2^2} \left(\frac{\delta}{\pi}\right) \{[\nu - w(u_1 \cos\Theta + u_2 \sin\Theta) - \nu_0]^2 + \delta^2\}^{-1} \quad (13\text{-}27)$$

Converting to dimensionless units, and again using the definition of the Voigt function, we find

$$R_{\text{III}}(x', \mathbf{n}'; x, \mathbf{n}) = \frac{g(\mathbf{n}', \mathbf{n})}{\pi^2} \sigma \int_{-\infty}^{\infty} \frac{e^{-u^2} H(\sigma, x \csc\Theta - u \operatorname{ctn}\Theta)}{(x' - u)^2 + a^2}\, du \quad (13\text{-}28)$$

where $\sigma \equiv a \csc\Theta$ and $a \equiv (\delta/w)$. This result is no longer expressible in terms of simple functions, and must be evaluated by numerical integrations (**529**).

Exercise 13-4: Fill in the steps required to obtain equation (13-28).

Exercise 13-5: In the limit that Compton scattering effects are negligible (i.e., $h\nu/mc^2 \ll 1$ where m is the mass of an electron), the scattering of radiation by electrons is coherent, $p(\xi', \xi) = \delta(\xi' - \xi)$, and grey, $f(\xi) = \sigma_e = (8\pi e^4/3m^2c^4)$. (a) Show that the redistribution function for electron scattering is

$$R_e(\nu', \mathbf{n}'; \nu, \mathbf{n}) = g(\mathbf{n}', \mathbf{n}) \left[\frac{mc^2}{4\pi kT(1 - \cos\Theta)\nu^2}\right]^{\frac{1}{2}} \exp\left[\frac{-mc^2(\nu - \nu')^2}{4kT(1 - \cos\Theta)\nu^2}\right]$$

[See (**196**; **475**; **161**; §86)]. (b) Show that for $T = 10^4$ °K, $\lambda = 4000$ Å, $\Theta = \pi/2$, the characteristic width of the above redistribution is about 10 Å. This result suggests that spectral lines in early-type stars can be significantly broadened by electron scattering (**159**; **475**; **161**, §86; **318**; **39**).

SYMMETRY PROPERTIES

The functions obtained above have certain symmetry properties which often may be exploited to simplify the form of the transfer equation. To obtain these, it is convenient to rewrite the general result (13-14) in terms of x and x'; let $\tilde{f}(\nu - \nu_0) \equiv f(\nu)$ and $\tilde{p}(\nu - \nu_0, \nu' - \nu_0) \equiv p(\nu, \nu')$. Then from equations (13-14), and (13-20) through (13-22)

$$R(x', \mathbf{n}'; x, \mathbf{n}) = w^2\pi^{-1} g(\mathbf{n}', \mathbf{n}) \int_{-\infty}^{\infty} du_1\, e^{-u_1^2} \int_{\infty}^{\infty} du_2$$

$$\times\, e^{-u_2^2} \tilde{f}[w(x' - \alpha u_1 - \beta u_2)] \tilde{p}[w(x' - \alpha u_1 - \beta u_2), w(x - \alpha u_1 + \beta u_2)] \quad (13\text{-}29)$$

Now for all four cases defined in §13-1, the scattering is symmetric in the atom's frame—i.e.,

$$\tilde{f}(\xi')\tilde{p}(\xi', \xi) = \tilde{f}(-\xi')\tilde{p}(-\xi', -\xi) \quad (13\text{-}30)$$

Suppose we replace $x' \to -x'$ and $x \to -x$ in equation (13-29); if we simultaneously change the signs on u_1 and u_2 [which we may do because the integrals span the entire range $(-\infty, \infty)$], we immediately see from equation (13-30) that

$$R(-x', \mathbf{n}'; -x, \mathbf{n}) = R(x', \mathbf{n}'; x, \mathbf{n}) \quad (13\text{-}31)$$

Next, noting that replacing $\mathbf{n}' \to -\mathbf{n}'$ and $\mathbf{n} \to -\mathbf{n}$ changes the sign of α and β, and recalling that $g(\mathbf{n}', \mathbf{n})$ depends only on $\mathbf{n}' \cdot \mathbf{n}$, it is clear from equations (13-29) and (13-30) that

$$R(-x', -\mathbf{n}'; -x, -\mathbf{n}) = R(x', \mathbf{n}'; x, \mathbf{n}) \quad (13\text{-}32)$$

and from (13-31) and (13-32) it follows that

$$R(x', -\mathbf{n}'; x, -\mathbf{n}) = R(x', \mathbf{n}'; x, \mathbf{n}) \quad (13\text{-}33)$$

Equations (13-31) through (13-33) depend only on the validity of (13-30) and hence are rather generally true. Further, because R depends only on the angle between $\mathbf{n}'$ and $\mathbf{n}$, we may interchange them without changing R; i.e.,

$$R(x', \mathbf{n}; x, \mathbf{n}') = R(x', \mathbf{n}'; x, \mathbf{n}) \quad (13\text{-}34)$$

Next we notice that, for coherent scattering in the atom's frame (Cases I and II), the redistribution functions depend only upon $(x + x')$ and $|x - x'|$ [cf. equations (13-23) and (13-26)]; hence x and x' may be interchanged without altering R; i.e.,

$$R_i(x, \mathbf{n}'; x', \mathbf{n}) = R_i(x', \mathbf{n}'; x, \mathbf{n}), \quad (i = \text{I, II}) \quad (13\text{-}35)$$

Finally, a transformation from $\mathbf{n}' \to -\mathbf{n}'$ changes the sign of α and β in the expressions for incoming photons, hence if $g(\mathbf{n}', \mathbf{n})$ is an even function of $\mathbf{n}' \cdot \mathbf{n}$,

$$R(-x', -\mathbf{n}'; x, \mathbf{n}) = w^2\pi^{-1} g(\mathbf{n}', \mathbf{n}) \int_{-\infty}^{\infty} du_1\, e^{-u_1^2} \int_{-\infty}^{\infty} du_2$$

$$\times\, e^{-u_2^2} \tilde{f}[w(-x' + \alpha u_1 + \beta u_2)] \tilde{p}[w(-x' + \alpha u_1 + \beta u_2), w(x - \alpha u_1 + \beta u_2)] \quad (13\text{-}36)$$

which will reduce to (13-29) in certain cases. In particular, the integrals will be equal if $\tilde{f}(\xi') = \delta(\xi')$ or if $\tilde{f}(\xi')\tilde{p}(\xi', \xi) = \tilde{f}(-\xi')\tilde{p}(-\xi', \xi)$. The former is true for Case I, while the latter is true for Case III where $\tilde{f}(\xi') = \tilde{f}(-\xi')$

and $\tilde{p}(\xi', \xi) = \tilde{f}(\xi)$. Hence we conclude that

$$R_i(-x', -\mathbf{n}'; x, \mathbf{n}) = R_i(-x', \mathbf{n}'; x, -\mathbf{n}) = R_i(x', \mathbf{n}'; x, \mathbf{n}), \quad (i = \text{I, III}) \quad (13\text{-}37)$$

where the second identity follows from equation (13-33).

APPLICATIONS

The solution of the transfer equation allowing for full angle and frequency redistribution is fairly difficult, and requires techniques more powerful than those discussed in this book; both Monte Carlo methods (**31**; **50**) and difference-equation methods (**460**) have been developed to handle the problem. There are a number of situations where both angular and frequency effects may become important. For example, the escape of $L\alpha$ photons from a very thick nebula depends sensitively upon the details of the scattering process, and a complete, careful treatment is required (**31**). In addition, when the material has a macroscopic velocity, both the absorption and emission coefficients become angle-dependent in the observer's rest frame (cf. §2-1). There then results an inextricable coupling between angles and frequencies in the transfer problem (see §14-1), and the details of the redistribution process may be of consequence; in this book, however, we shall consider only the case of complete redistribution in moving media, and will not pursue this point further. Finally, when observable angle-dependent information is available (e.g., center-to-limb variations as measured in the solar atmosphere), the possible importance of angular effects in the redistribution process must be examined. The investigations of this question conducted thus far (**640**; **641**; **460**) show that, at least for semi-infinite plane-parallel atmospheres with homogeneous layers, the differences between results obtained using angle-averaged redistribution functions and those obtained using the full angle-dependent functions are negligible. On the other hand, for optically thin media or small-scale structures (e.g., chromospheric fine-structure, spicules, etc.) the angular effects are often quite important (**186**), and should be considered in detail. However, these questions lie beyond the scope of this book, and will not be discussed here; rather, we now turn to the derivation of angle-averaged redistribution functions, and then to their use in radiative transfer calculations.

13-3 Angle-Averaged Redistribution Functions

As was noted in §2-1, if we assume that the radiation at any particular point in the atmosphere is essentially isotropic, then the contribution to the emissivity from scattering processes can be written in the form

$$\eta^S(\mathbf{r}, \nu) = \sigma(\mathbf{r}) \int_{-\infty}^{\infty} R(\nu', \nu) J_{\nu'} \, d\nu' \quad (13\text{-}38)$$

where $J_{\nu'}$ is the mean intensity and $R(\nu', \nu)$ is the *angle-averaged redistribution function*

$$R(\nu', \nu) \equiv (4\pi)^{-1} \oint R(\nu', \mathbf{n}'; \nu, \mathbf{n}) \, d\omega' \quad (13\text{-}39)$$

The rationale for this approach is essentially as follows. The radiation field will depart significantly from isotropy at any given frequency only at points whose optical depths τ_ν from the surface are of order unity or less; at depths $\tau_\nu \gtrsim 1$, the radiation field *is* essentially isotropic. Now a basic characteristic of non-LTE line-formation that emerged clearly in Chapter 11 is that the surface value of the source function is determined by photons contributed from over *an entire destruction length*, and over virtually all of this region the radiation field will, in fact, be isotropic. We may expect, therefore, that even at the surface where $I(\mu, \nu)$ shows departures from isotropy, S_ν will still have a value *already fixed* by processes occurring at depths where the anisotropy is negligible; hence the value of $I(\tau_\nu = 0, \mu, \nu)$ computed from this S_ν should be quite accurate. On the other hand, use of the angle-averaged functions accounts completely for the frequency-reshuffling of scattered photons, the action of which (as we saw in Chapter 11) affects crucially the photon escape-probability, and hence the thermalization process. In short, in this approach we account for the critical aspects of the redistribution process, and sacrifice information only in an area of secondary importance.

The functions defined by equation (13-39) are normalized such that

$$\int_{-\infty}^{\infty} d\nu \int_{-\infty}^{\infty} d\nu' \, R(\nu', \nu) = 1 \quad (13\text{-}40)$$

Integration over all emitted photons yields the absorption profile

$$d\nu' \int_{-\infty}^{\infty} R(\nu', \nu) \, d\nu = \phi(\nu') \, d\nu' \quad (13\text{-}41)$$

while integration over all absorptions yields the *natural-excitation emission profile* [cf. equation (2-14)]

$$\psi^*(\nu) \, d\nu = d\nu \int_{-\infty}^{\infty} R(\nu', \nu) \, d\nu' \quad (13\text{-}42)$$

If $R(\nu', \nu) = R(\nu, \nu')$ (we shall find this to be true for cases I–III), then clearly $\psi^*(\nu) = \phi(\nu)$. It should be emphasized, however, that $\psi^*(\nu)$ is *not*, in general,

the *actual emission profile* in the line when J_ν varies over the line profile (see the discussion in §13-4); *that* profile follows, ultimately, from the equations of statistical equilibrium *and* knowledge of the frequency-variation of J_ν.

Two hypothetical limiting cases sometimes considered are strict coherence and complete redistribution in the *laboratory frame*. In the former circumstance we would have

$$R(\nu', \nu) = \phi(\nu')\,\delta(\nu - \nu') \tag{13-43}$$

and in the latter

$$R(\nu', \nu) = \phi(\nu')\phi(\nu) \tag{13-44}$$

Neither of these limits can actually ever be achieved. As we have seen already, even if the scattering is strictly coherent in the atom's frame it is not in the laboratory frame unless the atoms have zero velocities (impossible). Further, as we shall see below, complete redistribution in the atom's frame does not produce exactly complete redistribution in the laboratory frame. However, in the latter situation it is nonetheless found that equation (13-44) does, in fact, provide a very good approximation to reality in most cases, and hence may normally be used to describe that limit. Note in passing that, because the dependences of R on ν' and on ν are *separated* in equation (13-44), the true emission profile $\psi_\nu \equiv \psi_\nu^* = \phi_\nu$ *regardless of the behavior of* J_ν, which justifies our earlier use of this relation [but only if equation (13-44) is actually accurate!].

GENERAL FORMULAE

The integration indicated in equation (13-39) could, in principle, be carried out directly, using each of the redistribution functions derived in §13-2. This however, turns out to be rather complicated. It is simpler to derive first a general formula, by performing the angle-average for arbitrary $f(\xi')$ and $p(\xi', \xi)$, and then to obtain specific forms for the particular cases of interest. We begin by rewriting equation (13-8), using the Doppler units defined in equations (13-11) and (13-12):

$$R_u(\nu', \mathbf{n}'; \nu, \mathbf{n}) = f(\nu' - w\mathbf{u}\cdot\mathbf{n}')p(\nu' - w\mathbf{u}\cdot\mathbf{n}', \nu - w\mathbf{u}\cdot\mathbf{n})g(\mathbf{n}', \mathbf{n}) \tag{13-45}$$

We now wish to fix ν' and ν and integrate over all angle. Choose an orthonormal triad $(\mathbf{n}_1, \mathbf{n}_2, \mathbf{n}_3)$ such that $\mathbf{u} = u\mathbf{n}_3$. Then $\mathbf{u}\cdot\mathbf{n} = \mu u$ and $\mathbf{u}\cdot\mathbf{n}' = \mu' u$, where $\mu \equiv \mathbf{n}\cdot\mathbf{n}_3$ and $\mu' \equiv \mathbf{n}'\cdot\mathbf{n}_3$. An element of solid angle may be written $d\omega = d\mu\, d\phi$ where ϕ is the azimuthal angle around $\mathbf{n}_3$. The phase function $g(\mathbf{n}', \mathbf{n})$ can be expressed in general as $g(\mu', \mu, \phi)$. Thus, angle-averaging

equation (13-45) we have

$$R_u(\nu', \nu) = (16\pi^2)^{-1}\int_0^{2\pi} d\phi \int_{-1}^{1} d\mu'\, f(\nu' - w\mu' u) \times \int_{-1}^{1} d\mu\, p(\nu' - w\mu' u, \nu - w\mu u)\int_0^{2\pi} d\phi'\, g(\mu', \mu, \phi') \tag{13-46}$$

Define

$$g(\mu', \mu) \equiv (4\pi)^{-1}\int_0^{2\pi} g(\mu', \mu, \phi')\, d\phi' \tag{13-47}$$

Then, noting that the integral over ϕ in equation (13-46) is trivial, we have

$$R_u(\nu', \nu) = \tfrac{1}{2}\int_{-1}^{1} d\mu'\, f(\nu' - w\mu' u)\int_{-1}^{1} d\mu\, g(\mu', \mu)p(\nu' - w\mu' u, \nu - w\mu u) \tag{13-48}$$

For simplicity, in what follows only *isotropic scattering* in the atom's frame will be considered; in the case $g(\mu', \mu) \equiv \frac{1}{2}$ [formulae for a dipole phase function are given in (**313**)]. Applying this restriction we have

$$R_{A,u}(\nu', \nu) = \tfrac{1}{4}\int_{-1}^{1} d\mu'\, f(\nu' - w\mu' u)\int_{-1}^{1} d\mu\, p(\nu' - w\mu' u, \nu - w\mu u) \tag{13-49}$$

Although equation (13-49) is general, it is not convenient in the case of coherent scattering in the atom's frame because of complications that arise in setting the limits of integration; let us, therefore, derive a more refined formula for this case.

If the scattering is coherent in the atomic frame,

$$p(\nu' - w\mu' u, \nu - w\mu u) = \delta[\nu' - \nu - wu(\mu' - \mu)] \tag{13-50}$$

Because the range of integration for μ' and μ is only $(-1, 1)$, it is clear that for a given value of u, the singularity of the δ-function will be outside the range of integration for sufficiently large values of $|\nu' - \nu|$, and $R_{A,u}(\nu', \nu)$ will, accordingly, be zero. Physically this corresponds to the fact that an atom moving with velocity u can change a photon's frequency by *no more than* $2uw$, this maximum shift occurring if the propagation vectors of the incoming and outgoing photons lie along the velocity vector and are oppositely directed. We substitute equation (13-50) into (13-49), and consider first the integration over μ. Let $y \equiv w\mu u$, and write

$$I = (wu)^{-1}\int_{-wu}^{wu} \delta[y - (\nu - \nu' + wu\mu')]\, dy \tag{13-51}$$

The integral will equal $(1/wu)$ if $-wu \leqslant \nu - \nu' + wu\mu' \leqslant wu$, and will be zero otherwise. Define $\Lambda(x)$ such that $\Lambda = 1$ if $-1 \leqslant x \leqslant 1$, and $\Lambda = 0$

otherwise. Then equation (13-49) can be rewritten, using equation (13-51), as

$$R_{A,u}(\nu', \nu) = (4wu)^{-1} \int_{-1}^{1} f(\nu' - wu\mu')\Lambda[\mu' + (wu)^{-1}(\nu - \nu')]\, d\mu' \quad (13\text{-}52)$$

If u is sufficiently small, then $|(\nu - \nu')/wu| > 1$ and Λ will vanish for all values of μ'. Thus there is a minimum speed $u_{\min}$, whose value we must determine, for which scattering from ν' to ν can actually occur. Define $\overline{\nu} \equiv \max(\nu', \nu)$ and $\underline{\nu} \equiv \min(\nu', \nu)$. First suppose that $\nu > \nu'$; then the requirement that the argument of the Λ-function fall in the range $(-1, 1)$ implies that $[(\nu - \nu')/wu] - 1 = [(\overline{\nu} - \underline{\nu})/wu] - 1 \leqslant 1$; this inequality yields

$$u_{\min} = (\overline{\nu} - \underline{\nu})/2w = |\nu - \nu'|/2w \quad (13\text{-}53)$$

But the same result is obtained by similar reasoning if we assume $\nu' > \nu$; therefore equation (13-53) is general. For $u \leqslant u_{\min}$, R_u will be zero. For $u > u_{\min}$, a contribution to R_u will come from part of the range of integration over μ'. To determine this range, we suppose, for definiteness, that $\nu > \nu'$. Then a contribution is obtained when

$$-1 \leqslant \mu' \leqslant 1 - [(\nu - \nu')/wu] = 1 - [(\overline{\nu} - \underline{\nu})/wu]$$

which implies that $\overline{\nu} - wu \leqslant \underline{\nu} - wu\mu' \leqslant \underline{\nu} + wu$; recalling that $\underline{\nu} = \nu'$ by hypothesis, the result just stated can be rewritten as

$$\overline{\nu} - wu \leqslant \nu' - wu\mu' \leqslant \underline{\nu} + wu \quad (13\text{-}54)$$

If we assume instead that $\nu' > \nu$, we again obtain equation (13-54), which is, therefore, general.

Now introducing the *Heaviside function* $\Phi(x, x_0)$, defined such that $\Phi = 1$ when $x > x_0$ and $\Phi = 0$ when $x < x_0$, making the substitution $y \equiv \nu' - wu\mu'$, and using equations (13-53) and (13-54), equation (13-52) may finally be written as

$$R_{A,u}(\nu', \nu) = (4w^2u^2)^{-1}\Phi(u - |\nu - \nu'|/2w, 0) \int_{\overline{\nu}-wu}^{\underline{\nu}+wu} f(y)\, dy \quad (13\text{-}55)$$

Finally, the results expressed in equations (13-49) and (13-55) must be averaged over the Maxwellian velocity distribution

$$P(u)\, du = \pi^{-\frac{3}{2}} e^{-u^2}(4\pi u^2)\, du \quad (13\text{-}56)$$

From equation (13-55) we then have for *coherence* in the atom's frame (Cases I and II)

$$R_A(\nu', \nu) = (\pi^{\frac{1}{2}}w^2)^{-1} \int_{u_{\min}}^{\infty} du\, e^{-u^2} \int_{\overline{\nu}-wu}^{\underline{\nu}+wu} f(y)\, dy \quad (13\text{-}57)$$

From equation (13-49) we have, for *noncoherence* in the atom's frame (e.g., Case III),

$$R_A(\nu', \nu) = \pi^{-\frac{1}{2}} \int_0^{\infty} du\, u^2 e^{-u^2} \int_{-1}^{1} d\mu'\, f(\nu' - w\mu' u) \times \int_{-1}^{1} d\mu\, p(\nu' - w\mu' u, \nu - w\mu u) \quad (13\text{-}58)$$

Finally it is convenient to use Doppler units [cf. equations (13-12), (13-20), and (13-21)] and write

$$R_A(x', x) = R_A(\nu', \nu)(d\nu'/dx')(d\nu/dx) = w^2 R_A(\nu', \nu) \quad (13\text{-}59)$$

RESULTS FOR SPECIFIC CASES

(a) Case I. Here $f(y) = \delta(y - \nu_0)$; hence the integral over y in equation (13-57) is nonzero only if $\underline{\nu} + wu \geqslant \nu_0 \geqslant \overline{\nu} - wu$. This implies that $u_{\min}$ now becomes effectively $u'_{\min} = \max(|x'|, |x|)$, which clearly satisfies the inequality $u'_{\min} \geqslant u_{\min}$ as given by equation (13-53). Then from equations (13-57) and (13-59)

$$R_{I,A}(x', x) = \pi^{-\frac{1}{2}} \int_{u_{\min}}^{\infty} e^{-u^2}\, du = \tfrac{1}{2}\,\mathrm{erfc}(u'_{\min}) \quad (13\text{-}60)$$

where the complimentary error function is defined as

$$\mathrm{erfc}(x) \equiv 2\pi^{-\frac{1}{2}} \int_x^{\infty} e^{-z^2}\, dz \quad (13\text{-}61)$$

Substituting for $u'_{\min}$,

$$R_{I,A}(x', x) = \tfrac{1}{2}\,\mathrm{erfc}[\max(|x|, |x'|)] \quad (13\text{-}62)$$

This redistribution function is easy to compute from well-known approximation formulae for erfc(x) (**4**, 299); asymptotic formulae and results for dipole scattering are given in (**313**).

A plot of $R_{I,A}(x', x)/\phi(x')$ is shown in Figure 13-2; the curves are labeled with the incoming photon frequency x', and give the probability for emission at frequency x, per absorption, as a function of x. We see that a photon absorbed at frequency x' will be emitted with equal probability at all x such that $-|x'| \leqslant x \leqslant |x'|$, and with exponentially decreasing probability beyond this range. It is easy to understand this result. The absorption can occur only when $\xi' \equiv 0$ in the atom's frame. Therefore atoms absorbing at frequency x' in the laboratory frame have a velocity of at least x' Doppler units. As the emissions occur with equal probability in all directions, the photons can be redistributed with equal probability over the entire range $\pm x'$.

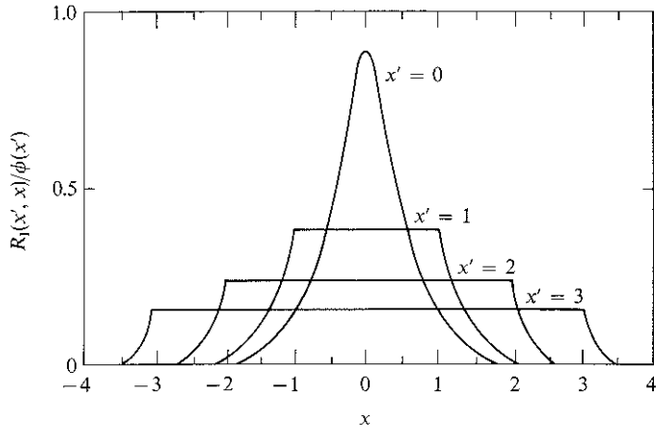

FIGURE 13-2
Probability of emission at frequency x, per absorption at frequency x', for pure Doppler redistribution. *Ordinate*: $R_I(x', x)/\phi(x')$; *abscissa*: emission frequency x. Curves are labeled with frequency x' of absorbed photon.

(b) Case II. Substituting equation (13-3) into (13-57), we have

$$R_{\text{II},\,A}(\nu', \nu) = (w^2\pi^{\frac{1}{2}})^{-1}\,\delta \int_{u_{\min}}^{\infty} du\, e^{-u^2} \int_{\bar{y}-wu}^{\underline{y}+wu} dy\,[(y - \nu_0)^2 + \delta^2]^{-1} \quad (13\text{-}63)$$

Converting to Doppler units, writing $z = (y - \nu_0)/\delta = x/a$, and writing $u_{\min} = \frac{1}{2}|x - x'|$ from equation (13-53) we have

$$\begin{aligned} R_{\text{II},\,A}(x', x) &= \pi^{-\frac{1}{2}} \int_{|x'-x|/2}^{\infty} du\, e^{-u^2} \int_{(\bar{x}-u)/a}^{(\underline{x}+u)/a} dz(1 + z^2)^{-1} \\ &= \pi^{-\frac{1}{2}} \int_{|x'-x|/2}^{\infty} e^{-u^2} \left[\tan^{-1}\left(\frac{\underline{x} + u}{a}\right) - \tan^{-1}\left(\frac{\bar{x} - u}{a}\right)\right] du \end{aligned} \quad (13\text{-}64)$$

where $\bar{x} \equiv \max(|x|, |x'|)$ and $\underline{x} \equiv \min(|x|, |x'|)$. The asymptotic behavior of R_{II}, and results for dipole scattering, are given in (**313**); an accurate method for the evaluation of R_{II} is given in (**6**).

The redistribution function R_{II} is of great interest, for it describes the important case of scattering by a resonance line that is broadened by radiation damping, and it has been extensively studied. A plot of $R_{\text{II},\,A}(x', x)/\phi(x')$ for $a = 10^{-3}$ is shown in Figure 13-3; again, the curves are labeled with the incoming photon frequency x', and give the probability, per absorption, of a subsequent emission at frequency x. Here we see that for small x' ($x' \lesssim 3$), the curves resemble those for R_{I}, because most of the

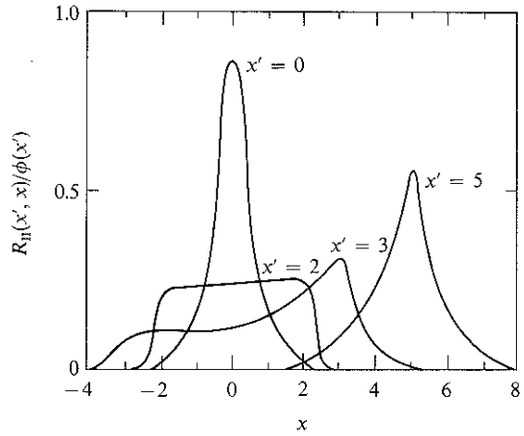

FIGURE 13-3
Probability of emission at frequency x, per absorption at frequency x', for Doppler redistribution from a coherently scattering (in the atom's frame) Lorentz profile with $a = 10^{-3}$. *Ordinate*: $R_{\text{II}}(x', x)/\phi(x')$; *abscissa*: emission frequency x. Curves are labeled with frequency x' of absorbed photon.

emissions are from atoms absorbing at line center and moving with velocities near x'. For large x' there will be very few atoms with high-enough velocity to absorb at line center, and most emissions then come from atoms moving with low velocities and absorbing in the line-wing. As the scattering process is coherent in the atom's frame, and the appropriate atoms are nearly at rest in the laboratory frame, the scattering will be nearly coherent in the laboratory frame as well. Thus *in the line-core* there is Doppler redistribution and *strong noncoherence*, while in the *wing* the scattering is more nearly *coherent*; this dichotomy has important implications for the line-transfer problem, as we shall see in §13-4.

In early work by Jefferies and White (**338**), it was suggested that a simple approximation for R_{II} could be written in the form

$$R_{\text{II},\,A}(x', x) \approx [1 - a(x)]\phi(x')\phi(x) + a(x)\phi(x)\,\delta(x - x') \quad (13\text{-}65)$$

where $a(x)$ is nearly zero for $x \lesssim 3$ and approximately unity for $x \gtrsim 3$. The original suggestion for the form of $a(x)$ is not adequate, however, and fails to meet requirements of normalization, symmetry, and wing-coherence. Nevertheless it is easy to define an appropriate function (**358**), and an approach of this type does simplify the calculation somewhat, although in the most precise work the correct form for R_{II} [equation (13-64)] should be used.

(c) Case III. In this case we no longer have coherence in the atom's frame, and equation (10-55) may be used without complication to give

$$R_{A,u}(\nu', \nu) = \frac{1}{4}\int_{-1}^{1} \frac{(\delta/\pi)\, d\mu'}{(\nu' - w\mu' u - \nu_0)^2 + \delta^2} \int_{-1}^{1} \frac{(\delta/\pi)\, d\mu}{(\nu - w\mu u - \nu_0)^2 + \delta^2}$$

$$= \frac{1}{4\pi^2 w^2 u^2}\left[\tan^{-1}\left(\frac{x' + u}{a}\right) - \tan^{-1}\left(\frac{x' - u}{a}\right)\right]$$

$$\times \left[\tan^{-1}\left(\frac{x + u}{a}\right) - \tan^{-1}\left(\frac{x - u}{a}\right)\right] \qquad (13\text{-}66)$$

Averaging over a Maxwellian velocity distribution and converting to Doppler units we have

$$R_{\mathrm{III},A}(x', x) = \pi^{-\frac{3}{2}} \int_0^\infty e^{-u^2}\left[\tan^{-1}\left(\frac{x' + u}{a}\right) - \tan^{-1}\left(\frac{x' - u}{a}\right)\right]$$

$$\times \left[\tan^{-1}\left(\frac{x + u}{a}\right) - \tan^{-1}\left(\frac{x - u}{a}\right)\right] du \qquad (13\text{-}67)$$

Asymptotic formulae, results for dipole scattering, and computational methods for evaluation of R_{III} are given in (**313**, **212**; **529**).

A plot of $R_{\mathrm{III}}(x', x)/\phi(x')$ for $a = 10^{-3}$ is shown in Figure 13-4. For small x', most absorptions are at line center by atoms moving with velocities near x', hence redistribution occurs with equal probability over the range $-x' \leqslant x \leqslant x'$, as was true for R_{I} and R_{II}. Again, for large x' most absorptions occur in the line wings of nearly stationary atoms, but now the emitted photons are completely redistributed over the absorption profile in the atom's frame, hence $R_{\mathrm{III}}(x', x)/\phi(x') \to \phi(x)$ for $x' \gg 1$. It has, on occasion, been argued on intuitive grounds that if the redistribution process is completely noncoherent in the atom's frame, and if this is combined with random Doppler motions, then the redistribution should be completely noncoherent in the observer's frame as well; this conclusion is false, however, as is clearly shown by Figure 13-5. In fact, Doppler motions introduce a correlation between incoming and outgoing frequencies near the line core, and the deviations from complete frequency redistribution in the laboratory frame can be large. Despite these deviations, it turns out (see §13-4) that the assumption $R(x', x) = \phi(x')\phi(x)$ produces line profiles quite similar to those obtained from the exact $R_{\mathrm{III}}(x', x)$, and in practice, the case of complete noncoherence in the atom's frame may be treated as complete noncoherence in the laboratory frame also, without serious errors.

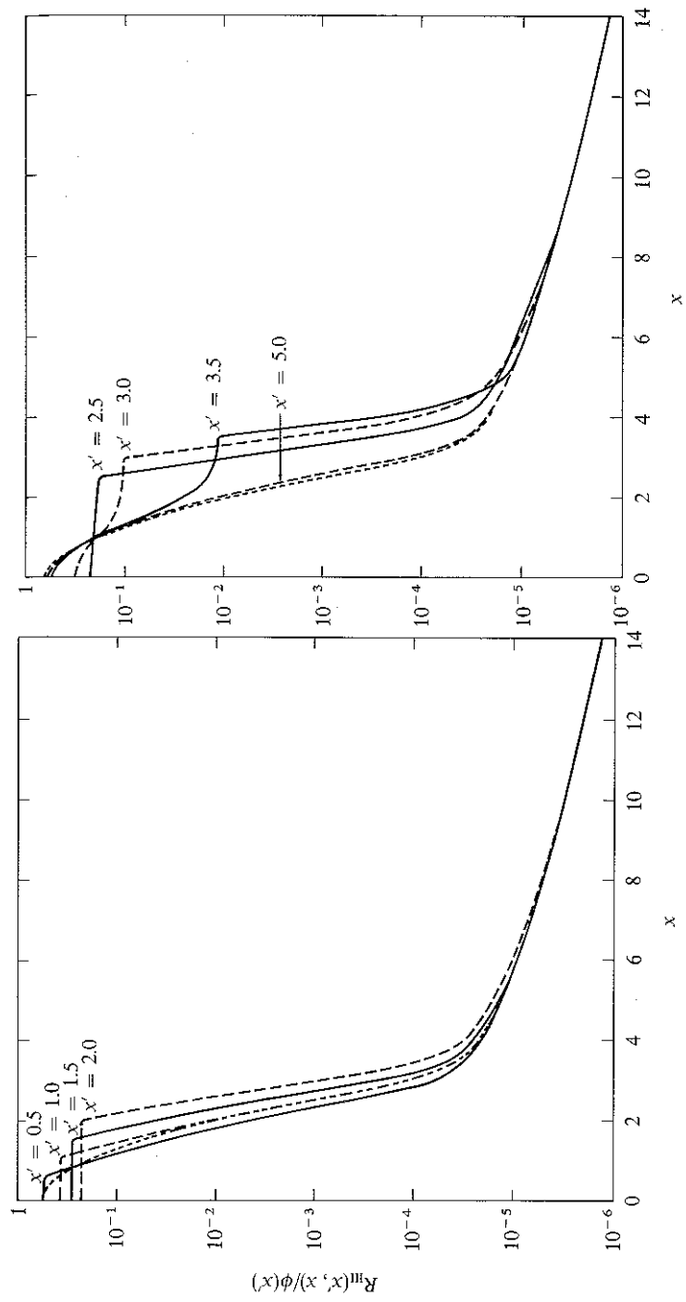

FIGURE 13-4
Probability of emission at frequency x, per absorption at frequency x', for complete redistribution over a Lorentz profile in the atom's frame, and Doppler redistribution in the laboratory frame, with $a = (\delta/w) = 10^{-3}$. *Ordinate:* $R_{\mathrm{III}}(x', x)/\phi(x')$; note logarithmic scale. Dotted curve is the natural-excitation emission profile $\phi(x)$. *Abscissa:* emission frequency x. Curves are labeled with frequency x' of absorbed photon. From (**212**), by permission.

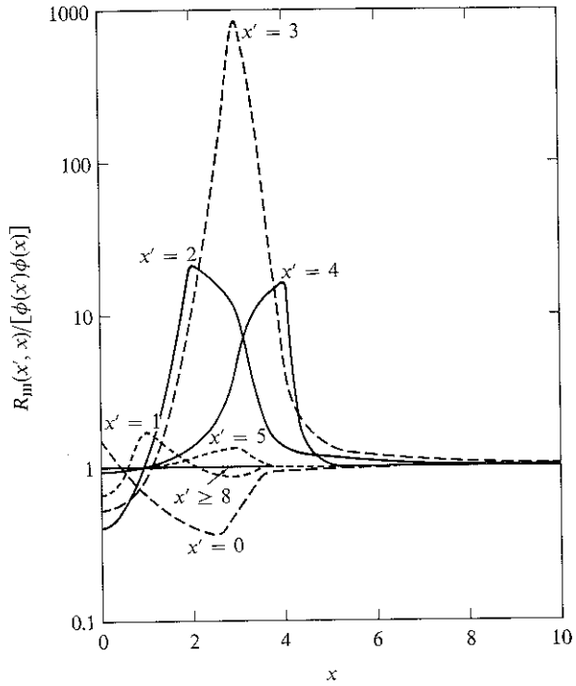

FIGURE 13-5
Ratio of actual redistribution function, for complete redistribution in the atom's frame and Doppler redistribution in the laboratory frame, $R_{III}(x', x)$, to the limiting case of complete noncoherence in the laboratory frame, $\phi(x')\phi(x)$. The absorption profile ϕ is a Voigt profile with $a = 10^{-3}$. From (**212**), by permission.

Exercise 13-6: Show that the angle-averaged redistribution function for scattering by electrons (under the same assumptions as employed in Exercise 13-5) is

$$R_{e,\,A}(\nu', \nu) = w^{-1}\,\text{ierfc}|(\nu' - \nu)/2w|$$

where $$\text{ierfc}(x) \equiv \int_x^\infty \text{erfc}(z)\,dz = \pi^{-\frac{1}{2}}e^{-x^2} - x\,\text{erfc}(x)$$

and w denotes the *electron* Doppler width, $w = \nu_0(2kT/m_e)^{\frac{1}{2}}/c$, which is about $43A^{\frac{1}{2}}$ times as large as the Doppler width of an atom of atomic weight A. See also (**318**) and (**39**).

SYMMETRY PROPERTIES

From equation (13-31) we see that, after angle-averaging, we must have

$$R(-x', -x) = R(x', x) \tag{13-68}$$

which is general so long as equation (13-30) holds. Further, by a direct extension of the arguments leading to equation (13-35) one finds

$$R_i(x', x) = R_i(x, x'), \quad (i = \text{I, II, III}) \tag{13-69}$$

The validity of this result can also be seen by inspection of equations (13-62), (13-64), and (13-67). From equation (13-69) it follows that the *natural-excitation* emission profile [see equation (13-42)]

$$\psi_i^*(x) \equiv \phi_i(x), \quad (i = \text{I, II, III}) \tag{13-70}$$

Exercise 13-7: Verify the result stated in equation (13-70) for Case I by direct integration of $R_{\text{I},\,A}(x', x)$ over x'.

13-4 Radiative Transfer with Partial Redistribution

Consider now the problem of accounting for the effects of angle-averaged partial redistribution in spectral line formation. For simplicity we shall confine attention to a two-level atom with continuum, and develop a general method that gives a rigorous solution of the problem. (This method can be generalized easily to more complicated atomic models as well.) Next, a somewhat simpler, but less general, approach will be developed, which still treats the troublesome stimulated emission term correctly. Then a still simpler method, which treats the stimulated emission terms only approximately, will be described, and results from the application of this method in highly idealized models will be discussed. Finally, results from calculations of resonance-line profiles, using the general method, for realistic model atoms and stellar atmospheres, will be presented.

FORMULATION FOR A TWO-LEVEL ATOM

Let us examine the formation of a resonance line connecting a perfectly sharp lower level to an upper state that is broadened into a distribution of substates by both radiation damping and collisions. Denote the population of the lower level by n_1, the total population of the upper level (summed over all substates) by n_2, and the number of ions by n_κ. The distribution of the atoms over the upper state is specified in terms of the *observer's-frame emission profile* ψ_ν, defined as the fraction of all atoms in the upper state that, if they decay radiatively, emit photons of frequency ν as seen in the laboratory frame. This method of counting atoms in the upper state is well-posed physically, for it describes the distribution in terms of *observables*; mathematically, ψ_ν is a complicated, but unique, one-to-one mapping from the distribution of atoms over their own rest-frame frequencies ξ, and velocities $\mathbf{v}$, the transformation being specified by the redistribution function

(**456**). Clearly ψ_ν is normalized, and $n_2 = n_2 \int \psi_\nu \, d\nu$. We shall find it convenient to define an auxilliary variable $\tilde{n}_2$ such that $n_2(\nu) = n_2\psi_\nu \equiv \tilde{n}_2(\nu)\phi_\nu$ gives the number of upper-state atoms in substates that can emit photons of frequency ν; the dominant variation of the upper-state distribution is factored out this way, and $\tilde{n}_2$ reflects only the departure from natural excitation.

The substate occupation number $n_2(\nu)$ (or, equivalently, ψ_ν) is specified by a rate equation of the form

$$n_2(\nu)(A_{21} + B_{21}J_\nu + C_{21} + R_{2\kappa} + C_{2\kappa}) = n_1\left[B_{12}\int R(\nu', \nu)J_{\nu'}\, d\nu' + C_{12}\psi_\nu^*\right] + n_2^*\psi_\nu^*(R_{\kappa 2} + C_{2\kappa}) \quad (13\text{-}71)$$

This equation has a simple physical interpretation [see also (**456**)]. The term on the lefthand side is the population of the substate, times the total rate of exit (a) to the ground state by spontaneous emissions, stimulated emissions and collisions, and (b) to the continuum by photoionizations and collisions. The only noteworthy term is that for stimulated emission which is *strictly coherent in both the laboratory and the atom's frame.* [This follows from the fundamental quantum-mechanical characteristics of the process (**197**, §62; **293**, §17), in which a photon, incident upon the atom, induces an emission in such a way as to create yet another photon of *exactly the same properties*: energy (hence frequency), momentum (hence direction of propagation), and polarization. Thus a photon $(\nu', \mathbf{n}')$ in the laboratory frame undergoes a definite transformation (depending on $\mathbf{n}'$ and $\mathbf{v}$) to $(\xi', \mathbf{n}')$ in the atom's frame, and creates another *identical* photon; both photons undergo *exactly* the same *inverse* transformation back into the laboratory frame, returning *two* photons $(\nu', \mathbf{n}')$ into the radiation field.] The terms on the righthand side describe processes of excitation to the substate $n_2(\nu)$, from the ground state and by recombinations from the continuum. The latter produce atoms in level 2 distributed according to the natural-excitation profile ψ_ν^*; n_2^* denotes the LTE particle density $n_2^* = n_\kappa n_e \Phi_2(T)$, where $\Phi(T)$ is the appropriate Saha–Boltzmann factor. For all three physically interesting cases of redistribution considered in §13-3, $\psi_\nu^* \equiv \phi_\nu$ [see equation (13-70)]. From the ground state, atoms excited by collisions are also distributed according to ψ_ν^*. The radiative excitations are given by the number of absorptions of photons of frequency ν', namely $n_1 B_{12} J_{\nu'}$, times the joint probability $R(\nu', \nu)$ of absorbing at ν' and emitting at ν, summed over all ν'. In the scattering process we assume that a fraction γ of all excited atoms emit coherently, and the remainder are completely redistributed in the atom's frame by elastic collisions [so that $p(\xi', \xi)$ is given by equation (13-6)]. The correct laboratory frame redistribution function in this case would be

$$R(\nu', \nu) = \gamma R_{\text{II}}(\nu', \nu) + (1 - \gamma)R_{\text{III}}(\nu', \nu) \quad (13\text{-}72)$$

but in practice it is adequate (and much easier) to use

$$R(\nu', \nu) = \gamma R_{\text{II}}(\nu', \nu) + (1 - \gamma)\phi_{\nu'}\phi_\nu \quad (13\text{-}73)$$

which assumes that complete redistribution in the atom's frame leads to complete redistribution in the observer's frame.

The rate equation for the ground state is

$$n_1\left(B_{12}\int \phi_\nu J_\nu \, d\nu + C_{12} + R_{1\kappa} + C_{1\kappa}\right) = n_2\left(A_{21} + B_{21}\int \psi_\nu J_\nu \, d\nu + C_{21}\right) + n_1^*(R_{\kappa 1} + C_{1\kappa}) \quad (13\text{-}74)$$

which has an interpretation entirely analogous to that of equation (13-71). Finally, the total number of atoms is presumed known, hence

$$n_1 + n_2 \int \psi_\nu \, d\nu + n_\kappa = n_{\text{atom}} \quad (13\text{-}75)$$

The required photoionization and recombination rates may be considered as given, in which event only the radiation field in the line needs to be calculated, or these rates may follow from a solution of the transfer equation in the continuum (assumed necessary for the ground state only).

The transfer equation to be solved may be written as

$$\partial^2(f_\nu J_\nu)/\partial\tau_\nu^2 = J_\nu - S_\nu \quad (13\text{-}76)$$

where, as usual, $d\tau_\nu \equiv -\chi_\nu \, dz$ and $S_\nu \equiv \eta_\nu/\chi_\nu$. In the line

$$\chi_\nu = \alpha_{12}[n_1 - (g_1/g_2)\tilde{n}_2(\nu)]\phi_\nu + X_\nu \quad (13\text{-}77)$$

and

$$\eta_\nu = (2h\nu^3/c^2)\alpha_{12}(g_1/g_2)\tilde{n}_2(\nu)\phi_\nu + E_\nu \quad (13\text{-}78)$$

and in the ground-state continuum

$$\chi_\nu = \alpha_{1\kappa}(\nu)(n_1 - n_1^* e^{-h\nu/kT}) + X_\nu \quad (13\text{-}79)$$

and

$$\eta_\nu = (2h\nu^3/c^2)e^{-h\nu/kT}\alpha_{1\kappa}(\nu)n_1^* + E_\nu \quad (13\text{-}80)$$

Here X_ν and E_ν represent (fixed) background opacity and emissivity sources, while $\alpha_{12} \equiv (B_{12}h\nu)/4\pi$.

The solution of the transfer equation is now quite complicated because the *emission profile ψ_ν is not known a priori*, but follows from the statistical equilibrium equations (13-71) through (13-75). Unlike the case of complete redistribution, where only the ratio (n_2/n_1) is required to specify the source function [cf. equation (11-4)] and hence only *one* statistical equilibrium equation is needed, we must now *compute* ψ_ν, and this introduces as many equations of the form of (13-71) as are required to define this function to

the desired precision. Indeed, the situation at hand strikingly resembles that of a multiline "multiplet" problem with very strong interlocking effects, each frequency within the line playing the role of a separate transition, and the line as a whole acting as the collective photon pool. This analogy suggests that the solution can be obtained effectively by means of a complete linearization technique.

METHODS OF SOLUTION

A general and powerful method of solving equations (13-71) through (13-75), simultaneously with the transfer equations (13-76) through (13-80), is to use the complete linearization technique. We introduce a discrete set of upper-state substates specified by frequencies $\{v_j\}$, measured relative to line center. The substate populations can then be written as $n_2(v_j) = \tilde{n}_{2j}\phi_j$, and the rate equations discretized as

$$n_1\left(B_{12}\sum_j w_j\phi_j J_j + C_{12} + R_{1\kappa} + C_{1\kappa}\right) - \sum_j w_j\tilde{n}_{2j}\phi_j(A_{21} + B_{21}J_j + C_{21}) - n_\kappa[n_e\Phi_1(T)(R_{\kappa 1} + C_{1\kappa})] = 0 \quad (13\text{-}81)$$

$$-n_1\left\{B_{12}\left[\gamma\sum_{j'}\mathscr{R}^{\mathrm{II}}_{jj'}J_{j'} + (1-\gamma)\phi_j\sum_{j'}w_{j'}\phi_{j'}J_{j'}\right] + C_{12}\phi_j\right\} + \tilde{n}_{2j}\phi_j(A_{21} + B_{21}J_j + C_{21} + R_{2\kappa} + C_{2\kappa}) - n_\kappa[n_e\Phi_2(T)(R_{\kappa 2} + C_{2\kappa})] = 0 \quad (13\text{-}82)$$

and

$$n_1 + \sum_j w_j\tilde{n}_{2j}\phi_j + n_\kappa = n_{\mathrm{atom}} \quad (13\text{-}83)$$

Here $\mathscr{R}^{\mathrm{II}}_{ij}$ is a discrete representation of $R^{\mathrm{II}}(v_i, v_j)$ in an appropriate quadrature. Writing $\mathbf{n} \equiv (n_1, \tilde{n}_{21}, \tilde{n}_{22}, \ldots, \tilde{n}_{2j}, \ldots, \tilde{n}_{2J}, n_\kappa)^T$, where J denotes the total number of substates, equations (13-81) through (13-83) are of the form $\mathscr{A}\mathbf{n} = \mathscr{B}$.

Suppose we have an estimate of the occupation numbers $\mathbf{n}$ for all depths; then current values of χ_ν, η_ν, and J_ν may be computed at all frequencies in the line and continuum. We may then linearize the transfer equation in terms of δJ_ν, $\delta\chi_\nu$, and $\delta\eta_\nu$, and express the latter two quantities in terms of δn_1, the δn_{2j}'s, and δn_κ. In turn, the δn's can be written as $\boldsymbol{\delta}\mathbf{n} = \sum(\partial\mathbf{n}/\partial J_k)\,\delta J_k$, where the sum extends over all frequencies in the line and continuum. The derivatives $(\partial\mathbf{n}/\partial J_k)$ can be written as $(\partial\mathbf{n}/\partial J_k) = -\mathscr{A}^{-1}[(\partial\mathscr{A}/\partial J_k)\cdot\mathbf{n}]$, and *explicit analytical* expressions can be obtained for the derivatives $(\partial\mathscr{A}/\partial J_k)$ [see (**456**; **459**)]. The final system of equations to be solved is of the standard Feautrier form $-\mathbf{A}_d\,\boldsymbol{\delta}\mathbf{J}_{d-1} + \mathbf{B}_d\,\boldsymbol{\delta}\mathbf{J}_d - \mathbf{C}_d\,\boldsymbol{\delta}\mathbf{J}_{d+1} = \mathbf{L}_d$ where

$$\boldsymbol{\delta}\mathbf{J}_d \equiv (\delta J_{1d}, \ldots, \delta J_{kd}, \ldots, \delta J_{Kd})^T.$$

When the system is solved, the δJ_ν's are applied to the current estimates of J_ν, and the rate equations (13-81) through (13-83) are re-solved for new $\mathbf{n}$'s. These values are used to compute χ_ν and η_ν, and a formal solution of the transfer equation updates the Eddington factors f_ν. The whole process is iterated to convergence. The formalism can be extended (**459**) to include the case of several lines from different sharp lower levels to common broadened upper levels [e.g., the Ca II H- and K-lines and infrared triplet, for which the $4s$ ground state and metastable $3d$ levels are sharp, while $4p$ is broad (see Figure 12-1)]. The convergence properties of this method are good, typically yielding a factor of 5 to 10 reduction in errors in the solution per iteration.

The method described above is effective, but is relatively expensive computationally, and it is worthwhile to explore less costly approaches (**292**). One of the basic problems encountered in treating partial redistribution is that the *unknown* emission profile appears explicitly in the stimulated emission term in χ_ν, and hence in the denominator of $S_\nu = \eta_\nu/\chi_\nu$. However, in many astrophysical applications, particularly for ultraviolet resonance lines in solar-type atmospheres, this stimulated emission correction, which is of order $\exp(-h\nu/kT)$, is extremely small. In this event we may proceed by *iterating* the ratio $\omega_\nu \equiv (\psi_\nu/\phi_\nu)$ in the stimulated emission term of an analytical expression for the source function obtained from manipulation of the statistical equilibrium equations.

Consider the source function for a strict two-level atom, omitting the continuum for simplicity. Using the Einstein relations we may write the line source function as

$$S_l(\nu) = \frac{(2h\nu^3/c^2)n_2\psi_\nu}{(g_2/g_1)n_1\phi_\nu - n_2\psi_\nu} = \frac{(2h\nu^3/c^2)(g_1n_2\psi_\nu/g_2n_1\phi_\nu)}{1 - (g_1n_2/g_2n_1)\omega_\nu} \quad (13\text{-}84)$$

Here, and in what follows, we assume that a current estimate of ω_ν is *known*. From equation (13-71), omitting the continuum terms,

$$(2h\nu^3/c^2)(g_1n_2\psi_\nu/g_2n_1\phi_\nu) = \frac{A_{21}\phi_\nu^{-1}\int R(\nu',\nu)J_{\nu'}\,d\nu' + C_{21}(1 - e^{-h\nu/kT})B_\nu(T)}{A_{21} + B_{21}J_\nu + C_{21}} \quad (13\text{-}85)$$

and from equation (13-74)

$$[1 - (g_1n_2/g_2n_1)\omega_\nu] = \frac{A_{21} + B_{21}(\bar{J}_e - \omega_\nu\bar{J}_a) + C_{21}(1 - \omega_\nu e^{-h\nu/kT})}{A_{21} + B_{21}\bar{J}_e + C_{21}} \quad (13\text{-}86)$$

where $\bar{J}_a \equiv \int\phi_\nu J_\nu\,d\nu$ and $\bar{J}_e \equiv \int\psi_\nu J_\nu\,d\nu = \int\phi_\nu\omega_\nu J_\nu\,d\nu$. Defining $\varepsilon' \equiv C_{21}(1 - e^{-h\nu/kT})/A_{21}$ we have

$$S_l(\nu) = \xi_\nu\left[\phi_\nu^{-1}\int R(\nu',\nu)J_{\nu'}\,d\nu' + \varepsilon'B_\nu\right] \quad (13\text{-}87)$$

where

$$\xi_\nu \equiv \frac{[(A_{21} + B_{21}\bar{J}_e + C_{21})/(A_{21} + B_{21}J_\nu + C_{21})]}{1 + (B_{21}/A_{21})(\bar{J}_e - \omega_\nu \bar{J}_a) + (C_{21}/A_{21})(1 - \omega_\nu e^{-h\nu/kT})} \tag{13-88}$$

Using equation (13-87) in an expression for the total source function of the form

$$S_\nu = [\chi_l(\nu)S_l(\nu) + \chi_c S_c]/[\chi_l(\nu) + \chi_c] \tag{13-89}$$

where now

$$\chi_l(\nu) = \alpha_{12}[n_1 - (g_1/g_2)n_2\omega_\nu]\phi_\nu + X_\nu \tag{13-90}$$

the transfer equation now reduces to the general form

$$\partial^2(f_\nu J_\nu)/\partial\tau_\nu^2 = J_\nu - a_\nu \int R(\nu', \nu)J_{\nu'}\, d\nu' + b_\nu \tag{13-91}$$

which may be solved straightaway by the standard Feautrier technique. In this method we use *current* estimates of ω_ν and J_ν to calculate ξ_ν in equation (13-88), and of the level populations and ω_ν to determine $\chi_l(\nu)$ in equations (13-89) and (13-90). The transfer equation (13-91) is then solved for improved values of J_ν. These J_ν's are then used in equations (13-81) through (13-83) to update n_1, $\tilde{n}_2(\nu)$, and n_κ, and from these, ψ_ν and thus ω_ν. The process is then iterated to convergence. One would expect this method to work well whenever $(B_{21}\bar{J}/A_{21}) \ll 1$; when stimulated emissions are very important one must use the full linearization technique described above.

RESULTS FROM IDEALIZED MODELS

A great deal of insight into the nature of partial redistribution effects can be obtained from studies similar in spirit to those described in Chapter 11, using idealized model atmospheres and atoms. To simplify the problem as much as possible, it is customary to make the additional physical assumption that the stimulated emission profile is given by ϕ_ν, not ψ_ν, and that the stimulated emission rate in equation (13-71) can be written as $n_2(\nu)B_{21}\bar{J}_a$ rather than $n_2(\nu)B_{21}J_\nu$. Then the parameter ω_ν defined above is identically unity, $\bar{J}_e \equiv \bar{J}_a$, and $\xi_\nu \equiv (1 + \varepsilon')^{-1}$, and equation (13-87) reduces to

$$S_l(\nu) = (1 - \varepsilon)\phi_\nu^{-1} \int R(\nu', \nu)J_{\nu'}\, d\nu' + \varepsilon B_\nu \tag{13-92}$$

where $\varepsilon \equiv \varepsilon'/(1 + \varepsilon')$. With this source function, the transfer equation can be solved in a single step without iteration. This approach has been used to estimate the differences between the frequency-dependent $S_l(\nu)$, obtained when partial redistribution effects are taken into account, and the frequency-independent

$$S_l^{\mathrm{CR}} = (1 - \varepsilon) \int \phi_\nu J_\nu\, d\nu + \varepsilon B_\nu \tag{13-93}$$

obtained from complete redistribution, for R_{I}, R_{II}, and R_{III} in constant-property atmospheres (**53**; **212**; **316**). The source functions $S_l(\nu)$ and S_l^{CR} may also be used to compute line-profiles; the errors made in calculated profiles if complete redistribution is assumed in place of an accurate redistribution function can then be assessed. An intermediate approximation is the *iterated* source function

$$S_l^{(1)}(\nu) \equiv (1 - \varepsilon)\phi_\nu^{-1} \int R(\nu', \nu)J_{\nu'}^{\mathrm{CR}}\, d\nu' + \varepsilon B_\nu \tag{13-94}$$

which is evaluated using the mean intensity obtained from the complete redistribution solution based on equation (13-93). It has been found that $S_l^{(1)}(\nu)$ in isothermal media is nearly equal to $S_l(\nu)$ (**316**), and obviously it is much simpler to compute.

Let us first consider the case of redistribution by Doppler shifts only—i.e., $R(\nu', \nu) = R_{\mathrm{I}}(\nu', \nu)$. A number of solutions for $S_l(\nu)$ have been obtained (**316**) for both finite and semi-infinite isothermal media, assuming zero continuum opacity, with $\varepsilon = 10^{-4}$ and 10^{-6}. Results for variation of the source function with frequency and depth are shown in Figure 13-6 for $\varepsilon = 10^{-4}$. The vertical arrows designate the frequency at which the monochromatic optical depth $\tau_x = 1$. It is clear that, at all frequencies where $\tau_x \gtrsim 1$, $S_l(x)$ essentially equals S_l^{CR}; large deviations occur when $\tau_x < 1$, but are of little consequence because these optically thin regions do not contribute significantly to the intensity in a line profile. In fact, the line profiles computed from S_l^{CR} and from $S_l(x)$ are virtually identical; hence we conclude that for Case I the scattering process is, for all practical purposes, adequately described by the simpler assumption of complete redistribution. A similar conclusion is reached (**212**) for the case of complete redistribution in the atom's frame—i.e., $R(\nu', \nu) = R_{\mathrm{III}}(\nu', \nu)$. Here the iterated source function, equation (13-94), was evaluated in an isothermal semi-infinite atmosphere, for lines with $a = 10^{-3}$, and $\varepsilon = 10^{-6}$ and 6×10^{-3}. Again it is found that $S_l^{(1)}(x)$ departs from S_l^{CR} only for $\tau_x < 1$, and that emergent line profiles computed from S_l^{CR} are almost identical to those computed from the frequency-dependent function. Thus the assumption of complete redistribution provides a very useful and accurate approximation for Case III as well.

The situation for Case II (coherent scattering in a broadened profile in the atom's frame with Doppler redistribution in the laboratory frame) is quite different. Results from an isothermal atmosphere of total thickness $T = 10^6$, for a line with $a = 10^{-3}$ and $\varepsilon = 10^{-4}$, given in (**316**), are shown in Figure 13-7. Here we see that at line center, $S_l(\nu)$ is near S_l^{CR} for shallow optical depths, but *rises above* S_l^{CR} at great depth, and *thermalizes* to the Planck function *sooner* than the complete-redistribution source function. This result is obtained because the coherent nature of the scattering process in the line-wings *inhibits* photon escape from the line-core, and forces more

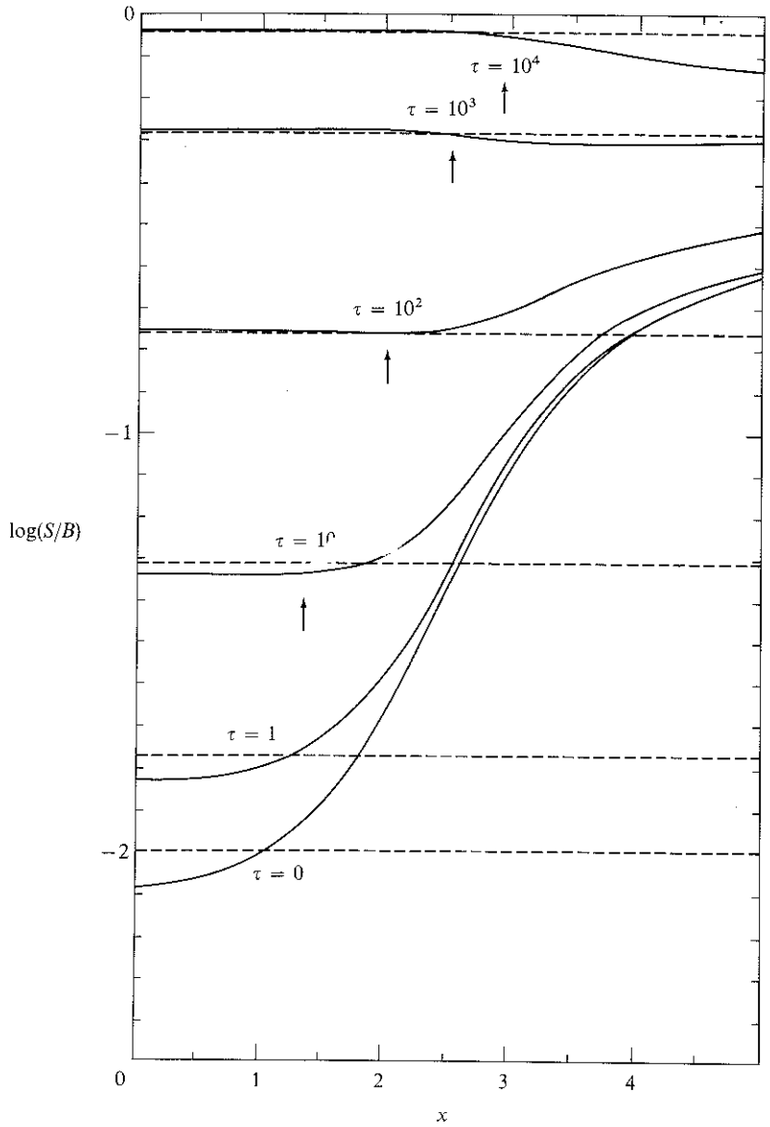

FIGURE 13-6
Source functions in an isothermal, semi-infinite atmosphere for a line with $\varepsilon = 10^{-4}$, assuming pure Doppler redistribution (case I). The vertical arrows show the frequency at which the monochromatic optical depth $\tau_x = 1$. *Dashed curves*: frequency-independent S_l^{CR} obtained assuming complete redistribution. *Solid curves*: frequency-dependent $S_l(x)$ obtained using correct redistribution function. *Abscissa*: displacement from line center in Doppler units. From (**316**), by permission.

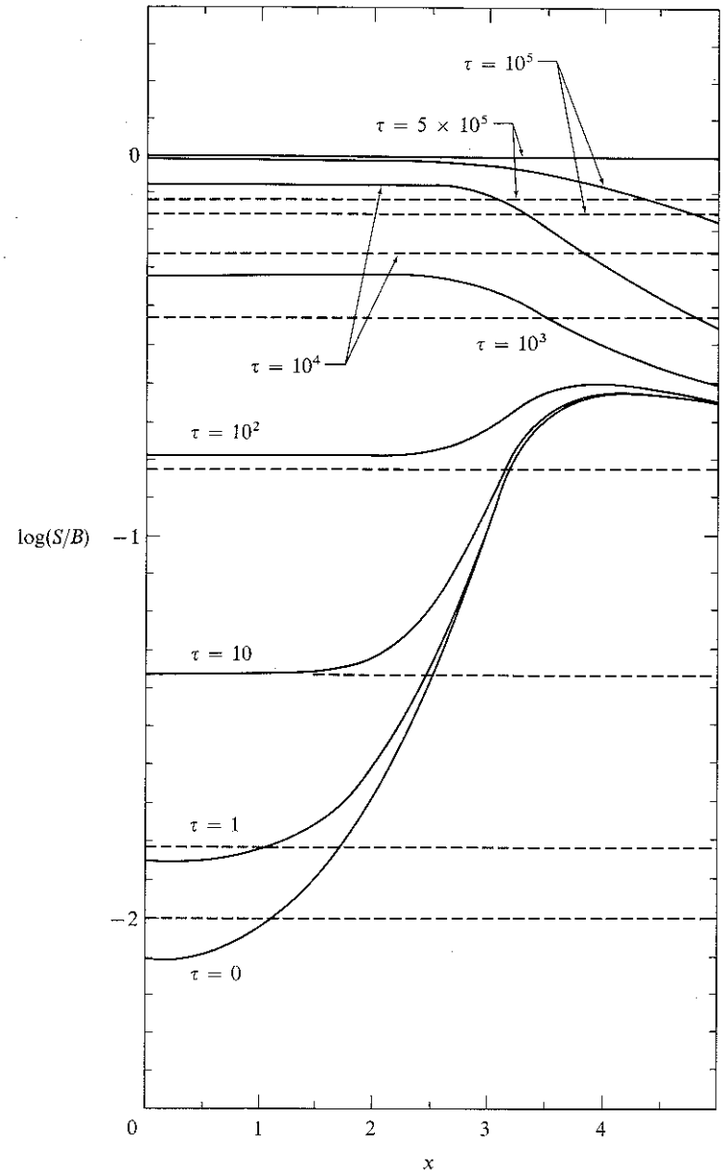

FIGURE 13-7
Source functions in an isothermal atmosphere of total thickness $T = 10^6$, for a line with $\varepsilon = 10^{-4}$ and $a = 10^{-3}$, assuming coherent scattering in a radiation-damping profile in the atom's frame, and Doppler redistribution in the laboratory frame. *Dashed curves*: frequency independent S_l^{CR} obtained assuming complete redistribution. *Solid curves*: frequency-dependent $S_l(x)$ obtained using correct redistribution function. *Abscissa*: frequency displacement from line-center in Doppler units. From (**316**), by permission.

rapid thermalization in the core by reducing the net escape probability for such photons. In contrast, in the line-wings, the frequency-dependent source function *lies substantially below* S_l^{CR}, because photons are no longer being fed into the wings from the core as efficiently. These effects are even more pronounced for finite atmospheres where S_l^{CR} may exceed $S_l(x)$ by orders of magnitude in the line wing! The emergent line profiles faithfully reflect the discrepancies between $S_l(x)$ and S_l^{CR}. Profiles for the correct source function have intensities that lie well below those for complete redistribution in the wings, and that nearly agree with the intensities predicted by coherent scattering. Although these effects will be diminished when there is an overlying continuum, we may, nevertheless, expect partial redistribution effects to be important for resonance lines formed in the outer layers of stars where densities, and thus collision rates, are low.

APPLICATION TO SOLAR AND STELLAR RESONANCE LINES

The resonance lines of many ions [e.g., those of hydrogen (particularly Lα), Ca^+, and Mg^+] are formed in chromospheric layers of low density, and hence are rather well characterized by the redistribution function of equation (13-73) with γ almost unity. Several calculations have now been made for solar lines, using the HSRA (**249**) or related models, and for solar-type giants (where densities are even lower than in the sun) using realistic model atmospheres.

One of the first examples of the importance of partial redistribution effects in resonance-line formation arose from attempts to fit the observed solar chromospheric Lα profile (**645**). It was found that when the line profile was computed under the assumption of complete redistribution, using models that provided accurate fits to the continuum data formed in the same atmospheric layers as the Lα line-wing, the intensity in the calculated profile wing was *much* larger (by a factor of 5 to 6) than observed. From the first study (**645**) it emerged that a much better fit to the profile is obtained if the scattering is assumed to be about 93 percent coherent, and only 7 percent completely redistributed. Subsequent work showed (**456**) that this parameterization is equivalent to using equation (13-73) with realistic values of γ (determined from the known collisional and radiative rates), and finally, that when the full depth-dependence of the profile and redistribution functions, atomic rates, level populations, and background opacity are taken into consideration, an excellent fit to the observations is achieved (**457**).

An even more interesting example is provided by the solar Ca II *H*- and *K*-lines, for which the earlier work (e.g., **401**) assuming complete redistribution gives a good fit to the disk-center profiles, but fails to fit the observed center-to-limb variation. Calculations using a five-level atom similar to that shown in Figure 12-1 were made (**570**) [see also (**642**)] for three model atmospheres: (1) the HSRA with a depth-independent microturbulence of

4.5 km s^{-1}; (2) the HSRA with the microturbulent velocity distribution given by (**401**); and (3) the same as model 1, but with the temperature structure modified to be $T_3 = \max(T_{HSRA}, 4450°K)$, which raises the temperature minimum by about 300°K. In all cases the full depth-variations of the line profile (allowing for radiation, van der Waals, and Stark broadening), the redistribution function, and background sources were taken into account.

The results from complete redistribution (CR) and partial redistribution (PR) computations of the double-reversal near line center (which provides vital information for diagnostics of the temperature-minimum region and the chromosphere) are strikingly different. As may be seen in Figure 13-8,

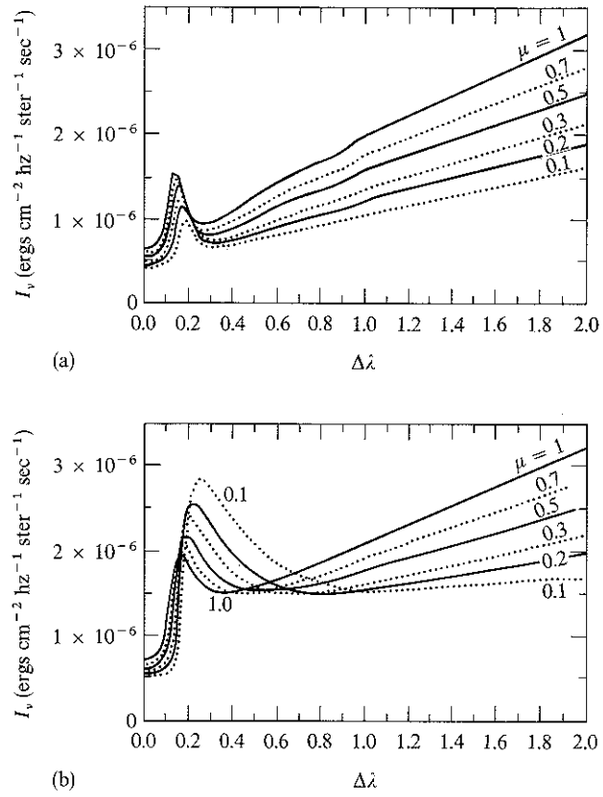

FIGURE 13-8
Ca II *K*-line profiles computed using the HSRA (**249**) model solar atmosphere with the distribution of microturbulent velocities given in (**401**). *Ordinate*: specific intensity $I(\mu, \Delta\lambda)$ in absolute units; *abscissa*: displacement $\Delta\lambda$ from line-center in Å. (a) Partial redistribution results; (b) complete redistribution results. Curves are labeled with μ, the cosine of the angle from disk center. From (**570**), by permission.

the PR profiles show uniform limb-darkening throughout the entire profile. In contrast, the CR results show limb-brightening at K_2 (the emission maximum) and no center-to-limb variation at K_1 (the minimum outside K_2); both of these results are contrary to observation. Furthermore, the wavelength position of the K_1 minimum shows a rapid center-to-limb increase for CR (again contrary to observation), while for PR the increase is much slower. The problem of limb-brightening of K_2 when CR is assumed can be overcome if a very special distribution of microturbulent velocities (**27**) is employed. However, neither of the problems just described for K_1 is eliminated in this way; moreover, the need for this special assumption is obviated when PR is employed.

The quantitative improvement in the comparison with the observed behavior of the K_1 feature is shown in Figure 13-9. There we see that CR produces much too rapid a rise in $\Delta\lambda(K_1)$ as $\mu \to 0$, while both models 1 and 3 yield an excellent fit to the data. Model 2 does not do as well, though it provides a better fit to certain data for K_2 (not discussed here). In part (b) of the figure we see that CR fails badly to fit the limb-darkening of K_1 intensity, while PR (with any of the three models) yields the correct center-to-limb *variation* (notice the *logarithmic* scale). Note that the *absolute intensity* of the K_1 feature is reduced by PR, relative to CR, for a *given model* (compare the open and filled circles at $\mu = 1$). This results from the essentially coherent nature of the scattering process in the K_1 feature, and is consistent with the results for R_{II} redistribution in the idealized models discussed above. One way to obtain a fit to the absolute intensity is to raise the temperature minimum by 300°K to 400°K over the HSRA value, as was done for model 3. While this change seemingly is small, it should be recalled that a change of about 100°K near T_{min} changes the energy content of that region by an amount comparable to the energy content of the entire corona (cf. §7-7.). One hesitates to conclude on the basis of one line that T_{min} must be higher than the HSRA model value; however, similar results are also obtained from an analysis of the Mg II *h*- and *k*-lines (**59**), and it may, in fact, be necessary to raise the empirical value of T_{min} to about 4400°K. Such an adjustment would bring the semiempirical result into harmony with the estimates of T_{min} from radiative equilibrium models, and would eliminate the difficulty described in §7-7. In any case it must be emphasized that the drop in intensity of PR relative to CR is a *differential effect for a given model*, and therefore will occur no matter what model is used. Thus if T_{min} is estimated from a fit to the K_1 intensity (as is sometimes done in stellar work), one *must* use a PR description of the scattering process, especially for giants (**571**), or *systematic* errors will be made. Finally, it is found that the PR calculation accurately reproduces the relative behavior of the *H*- and *K*-lines while CR does not.

The marked differences in the CR and PR predictions of limb-darkening and wavelength-position of K_1 can be easily understood in terms of the differences in the depth-variation of the source function S_l^{CR}, which is

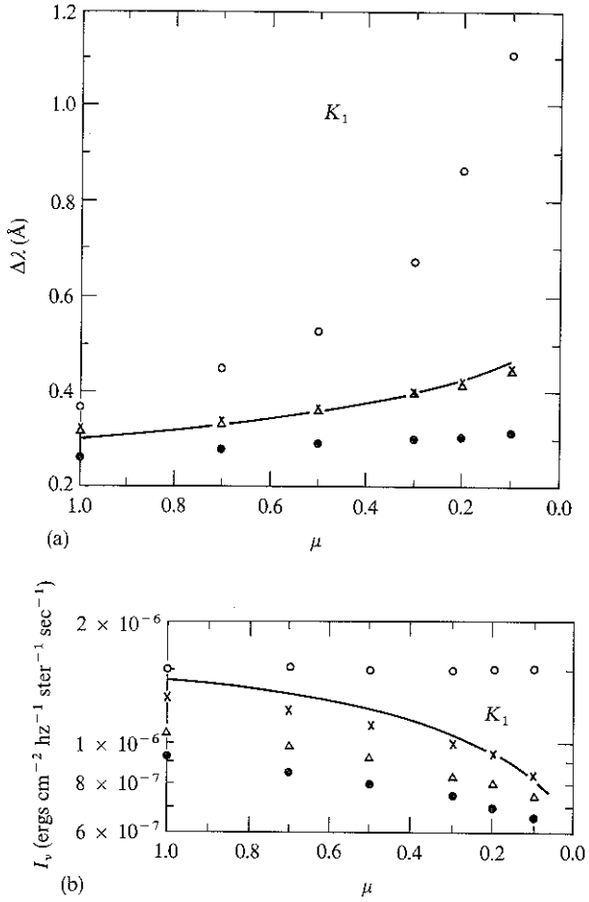

FIGURE 13-9
(a) Wavelength position of minimum intensity outside emission core, K_1, as a function of μ, cosine of angle from disk center. (b) Limb-darkening at K_1; note logarithmic units of intensity. *Open circles*: complete redistribution results, model 2 (see text). *Filled circles*: partial redistribution, model 2. *Open triangles*: partial redistribution, model 1 (see text). *Crosses*: partial redistribution, model 3 (see text). From (**570**), by permission.

frequency-independent, and the frequency-dependent PR source functions $S_l(\nu)$ shown in Figure 13-10. There we see that CR yields a unique source function that has a single absolute minimum. As one observes from center to limb, the slant-length optical depth at a particular frequency increases; if S_l^{CR} is used, this implies that the intensity at $\Delta\lambda_{K_1}(\mu = 1)$ *must rise* as μ decreases. The minimum S_l^{CR} will manifest itself only at some larger $\Delta\lambda$ where the material is more transparent and thus reflects the source function

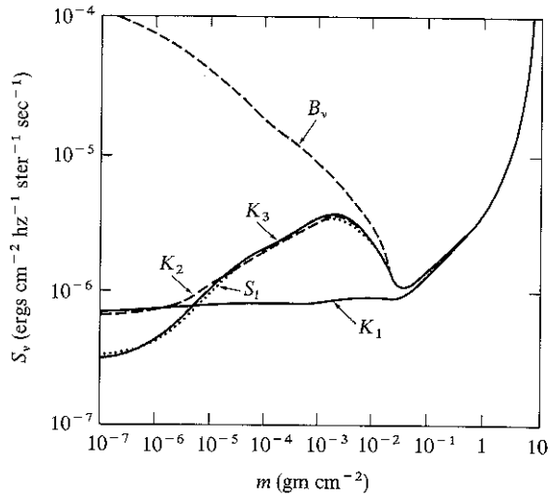

FIGURE 13-10
Depth-variation of K-line source function (for a five-level atom) and Planck function in model 1. *Dashed curve*: Planck function B_ν. *Dotted curve*: complete-redistribution source-function S_l (see also Figure 12-2). The other curves show the partial-redistribution source-functions at line-center (K_3), the emission peak (K_2), and the profile minimum outside the peak (K_1). *Abscissa*: column density gm/cm^2. From (**570**), by permission.

deeper in, near the minimum. This explains both the rapid increase in $\Delta\lambda(K_1)$ as a function of μ, and the near-constancy of I_{K_1} predicted by CR. In contrast, the PR source function at line-center (K_3) lies *above* the CR value (as a result of photon trapping as discussed previously). But at K_1 the source function has an essentially coherent-scattering character, is decoupled from other frequencies, and shows a monotonic decrease outward in the atmosphere [similar results were obtained in (**642**)]. For this behavior of S_l, changing the point of observation from center to limb merely samples $S_l(K_1)$ higher in the atmosphere, where it has (for PR) a *lower* value, and hence produces a decrease in I_{K_1}, as desired, at about the same $\Delta\lambda$. In short, taking partial redistribution into account yields a *substantial* improvement in the degree of agreement between theory and observation for the solar Ca II *H*- and *K*-lines. Similar calculations, yielding similar results, have also been made for the Mg II *h* and *k* (3s–3p) lines for the solar atmosphere (**458**; **59**) and for solar-type stars (**455**). It now appears that it is mandatory to account for partial redistribution effects in the interpretation of strong chromospheric resonance lines with strong radiation-damping wings, and further efforts in this direction will undoubtedly be amply rewarded.

14

Radiative Transfer in Moving Atmospheres

The existence of macroscopic motions (i.e., nonthermal velocities that are coherent over distances much larger than a particle mean-free-path) in stellar atmospheres is well-documented by a wealth of observational evidence. These motions appear to be present on all scales, from "eddies" whose sizes are small compared to a photon mean-free-path, up to expansion of the atmosphere as a whole. Although velocity fields have but little effect on radiative transfer in the continuum, they strongly influence line-formation because even a small (Doppler) frequency shift of a line produces a major change in its absorptivity as seen by a stationary observer.

In the analysis of spectra of supergiants, Struve and Elvey (**617**) discovered that the *Doppler widths inferred from the position of the flat part of the curve of growth* (cf. §§10-3 and 10-4) *were far in excess of the thermal value*. They attributed this broadening to nonthermal "turbulent" velocities, presumed to have a Gaussian distribution. The geometric scale of these motions (which are called "*microturbulence*") is supposed to be so small that they act as additional line-broadening agents, and thus *enhance the line-strength*. The inferred velocities often approach or *exceed* the speed of sound in the material, and it is clear that astrophysical "microturbulent" velocities are not to be identified

with turbulence in the strict fluid-dynamical sense, but rather with unresolved motions.

The picture is clarified when we examine the solar spectrum. From curves of growth, or from spectra of low spatial, spectral, and temporal resolution, one again infers a significant microturbulent velocity. But high-resolution spectra have a characteristic "wiggly-line" appearance [see Figure III-2 of (**20**) for an excellent example], showing distinct Doppler shifts and asymmetries that fluctuate rapidly along the length of the slit, and in time. Probably a large part, or even all, of the velocity field is composed of *wave motions* of various types, scales, and periods, which (when superposed) give rise to a pattern that appears chaotic and "turbulent". However, the diagnostics are still in a primitive state, and the precise nature of the velocity field is not at all well known.

Evidence for velocity patterns on a *large* scale ("*macroturbulence*") was provided by Struve's observation (**616**) that the *widths of line profiles* in certain stars exceeded the Doppler widths obtained from the curve of growth of their spectra; here the line strength is unchanged, and one envisions areas on the stellar surface so large as to be practically independent "atmospheres" moving systematically along the line of sight. A discussion of the observations, with numerous references, can be found in (**261**, Chap. 8). Further, periodic Doppler shifts of the lines in some stellar spectra reveal that they are from *pulsating* stars. Beyond this, objects such as the WR stars, P-Cygni stars, and early-type supergiants all show characteristic line profiles, with blue-shifted absorption components and red-shifted emission components, indicative of large-scale expansion.

It is clear that a characterization of stellar velocity fields by the two extremes of "micro"- and "macro"-turbulence is an oversimplification and, in the end, we wish to know *distribution functions* describing the amplitudes and scales of the velocity patterns. Further, one would like to relate observed parameters to more fundamental quantities such as the velocities of convective motions. Finally, a truly consistent theory of stellar atmospheres will require a *dynamical* theory of the interaction of material velocities, the thermodynamic state of the matter, and the radiation field; only then will we be able fully to understand stellar chromospheres and coronae.

At the present time, however, we are far from having such a complete theoretical structure. In this chapter we shall focus almost entirely on the "kinematics" of radiative transfer in moving media; i.e., *given* the velocity field and the model atmosphere, *compute* the emergent spectrum. A variety of techniques exist to attack the problem just posed. *Observer's-frame methods* can handle complicated velocity fields and multidimensional structures (although we shall confine attention to one-dimensional problems only), but are generally restricted to velocities of the order of a few Doppler widths, and hence are not well suited to handle rapid atmospheric expansion. *Sobolev's method*, on the other hand, provides an approximate solution in the case of rapid flow with large velocity gradients. *Comoving-frame methods* span the two extremes and provide general solutions applicable in both limits. Special techniques have been developed to treat *random* or *stochastic velocity fields.*

14-1 The Transfer Equation in the Observer's Frame

FORMULATION AND SOLUTION OF THE TRANSFER EQUATION

When the material in the atmosphere moves with velocity $\mathbf{v}(\mathbf{r})$ relative to an external observer at rest, there is a *Doppler shift* of photon frequencies between the observer's frame, and the frame of the atoms of which the material is composed. If the frequency in the observer's frame is ν, then in the atom's frame the frequency at which a photon traveling in direction $\mathbf{n}$ was emitted, or can be absorbed, is

$$\nu' = \nu - \nu_0(\mathbf{n} \cdot \mathbf{v}/c) \tag{14-1}$$

Thus, the opacity and emissivity of the material, as seen by a stationary observer, become *angle-dependent*. The transfer equations for a time-independent moving medium in planar geometry is then

$$\mu[\partial I(z, \mu, \nu)/\partial z] = \eta(z, \mu, \nu) - \chi(z, \mu, \nu)I(z, \mu, \nu) \tag{14-2}$$

It is convenient to measure frequency displacements from line center in units of a fiducial Doppler width $\Delta\nu_D^* \equiv \nu_0 v_{th}^*/c$, where v_{th}^* is a thermal velocity parameter, and to measure velocities in the same units, $V = v/v_{th}^*$. Then the transformation between observer's frame and atom's frame frequencies is

$$x' = x - \mu V \tag{14-3}$$

where $x \equiv (\nu - \nu_0)/\Delta\nu_D^*$ and x' is defined similarly. The effects of Doppler shifts are inconsequential for continuum terms, which do not vary much over the frequency range implied by velocity shifts, so we account only for changes in line terms and write

$$\chi(z, \mu, x) = \chi_c(z) + \chi_l(z)\phi(z, \mu, x) \tag{14-4}$$

and

$$\eta(z, \mu, x) = \eta_c(z) + \eta_l(z)\phi(z, \mu, x) \tag{14-5}$$

where the normalized line-profile is defined by

$$\phi(z, \mu, x) \equiv \phi(z; x - \mu V) \tag{14-6}$$

For example, for a Doppler profile,

$$\phi(z, \mu, x) = \pi^{-\frac{1}{2}}\,\delta^{-1}(z)\exp\{-[x - \mu V(z)]^2/\delta^2(z)\} \tag{14-7}$$