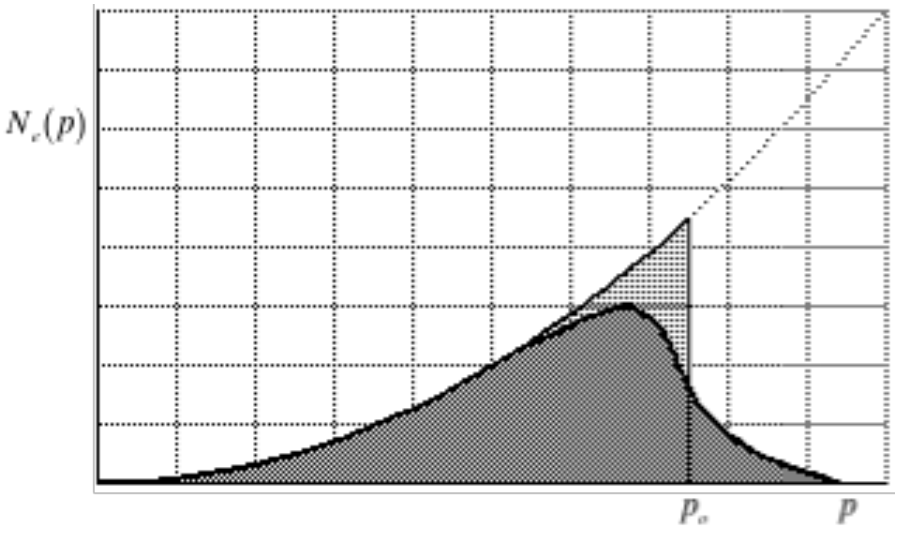
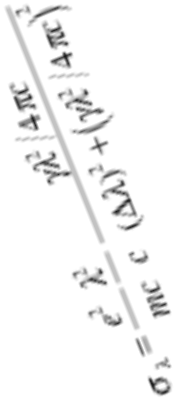
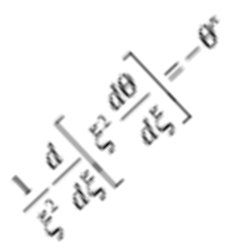
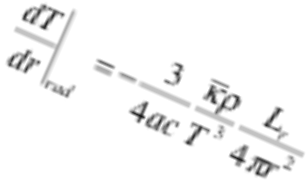




Notes on Stellar Astrophysics

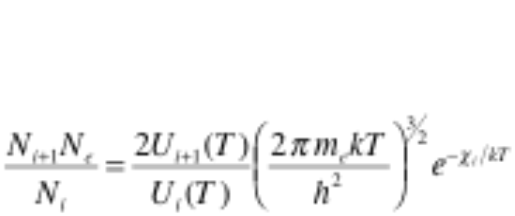


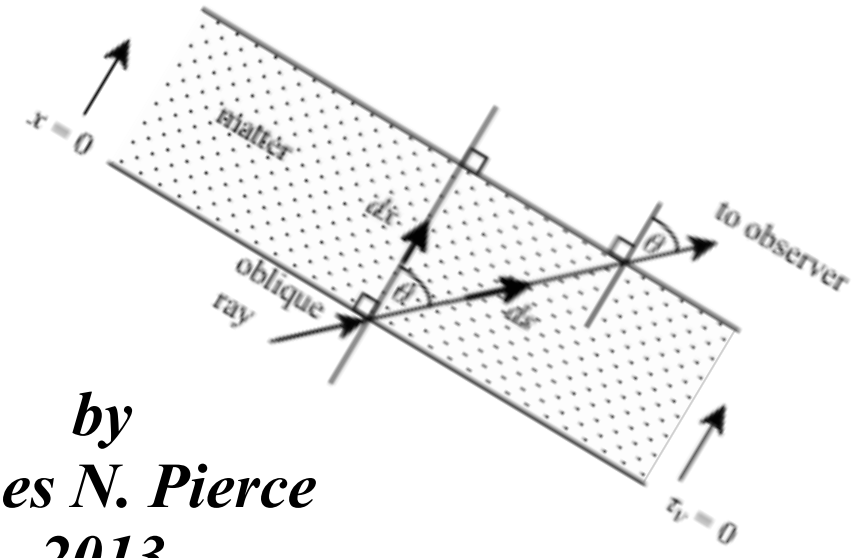


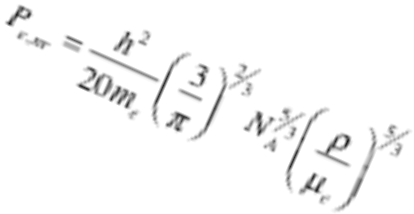
by

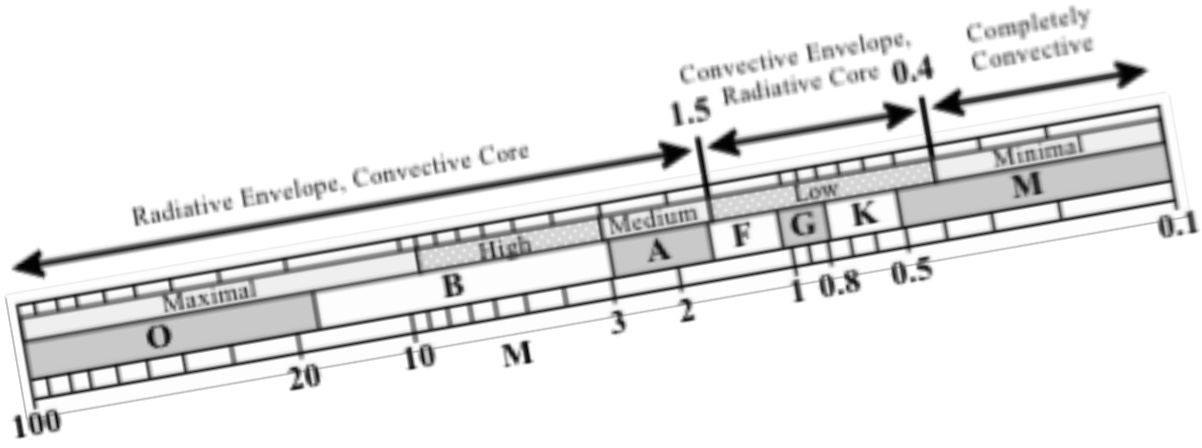
James N. Pierce

2013









To Lee Anne Willson and George H. Bowen

Edition 4.0

TABLE OF CONTENTS

PREFACE

This book is based primarily on lecture notes for a two-semester sequence of courses at Minnesota State University, Mankato: *Stellar Astrophysics* and *Stellar Structure and Evolution*. The courses, which were offered every other year, were taken by astronomy majors during their junior or senior years; these students were expected to have completed the three-semester *Calculus* sequence, *Differential Equations*, the *General Physics* sequence, the first semester of *Modern Physics*, and the sophomore-level *Astronomy and Astrophysics* sequence. Depending on the timing, students may or may not have been exposed to upper-level courses in *Mechanics*, *Electricity and Magnetism*, and *Quantum Mechanics*. Thus, some of this material would have been seen before, but normally not in the depth presented here. This course sequence was intended as part of the student's preparation for graduate study in astronomy or astrophysics, and it was expected that much of the material would be covered again at that time.

The focus of the book is on *understanding* the material in the courses, rather than just memorizing it. Wherever possible, derivations of equations are presented, with sufficient intermediate steps to lead the reader safely through the mathematical maze. In this way, the student will comprehend why and when approximations have been made and/or certain terms or factors have been ignored; this is important because it is necessary to understand the limits of the equations being developed in order to properly apply them. Additionally, it is useful for the student to be able to adapt existing equations to solve new problems.

Most of the material contained in this book has been covered by many other authors; this is *not* a completely new set of basic principles, equations, derivations, and results for use in understanding stars. Rather, it is a somewhat different organization of the standard collection of material, based on the order in which it is presented in the two courses involved. Neither an atmospheres/interiors split nor a principles/applications division has been used to allocate material to the two courses or to partition the book; instead, the approach is to develop and utilize topics as needed, such that the text flows smoothly from chapter to chapter in a logical order. In my sequence of courses, the first semester covers Chapters 1 - 7, and the second covers Chapters 8 - 15.

As this is a course text, an effort has been made to keep the book as readable as possible for the student; as a result, there are relatively few references included. Where appropriate, references are used to guide the reader to particular original sources – constants, data, models, etc. – but for the most part the story is not liberally referenced as most of the topics have already

been discussed in many other sources. A list of references is provided at the end of the book, along with a reading list of books on the subject, for those needing another slant on the story.

Many of the books used as references have been out of print for some time (which has provided much of the impetus for this project), but they are still good sources of material, including numerous graphs, tables, and models. A number of books from the reading list have been used as texts for these courses over the last 30 years or so, but finding a book with satisfactory content that is written at an appropriate level and is still in print has proven to be somewhat elusive of late. It is hoped that this text will serve as an adequate guide to basic stellar astrophysics for some time to come.

CHAPTER 1: Overview

To many astronomers, stars are points of light in the sky; these points may have measurable positions, velocities, magnitudes, colors, etc., but they are still just points. To the astrophysicist, however, stars are huge balls of matter, with properties such as temperature, density, pressure, composition, etc. that vary throughout the star and over time. This book provides an introduction to the details of the structure, operation, and evolution of stars.

Most stars are immense balls of gaseous matter; these gases are discouraged from simply drifting away into space by the star's gravity, which attracts the matter toward the center of the star. At the same time, they are prevented from all piling up at the center by the pressure within the star, which is supplied by both the matter and the radiation within the star. This pressure is greatest at the center of the star, and we can say that the pressure gradient is generally negative.

Eq. 1.1 $$\frac{dP}{dr} < 0$$

We shall find that the high pressure in a star is normally produced as a result of high temperature, and thus the temperature gradient should also be generally negative.

Eq. 1.2 $$\frac{dT}{dr} < 0$$

The high temperatures within a star serve several purposes:

- They maintain the matter in a gaseous state, usually a plasma of ions – nuclei and electrons.
- They produce radiation that ultimately escapes from the stellar surface as the star's luminosity.
- They result in high particle velocities that smash nuclei together in the core of the star to perform nuclear fusion, releasing energy that migrates outward towards the surface, maintaining the high temperatures along the way.

In studying stars, astrophysicists have found it convenient to draw boundaries that divide a star into different regions. The simplest boundary to understand is the **stellar surface**: everything inside the surface is part of the **stellar interior**, while the stellar gases that lie above the surface make up the **stellar atmosphere**. The gases of the stellar atmosphere are reasonably transparent, or optically thin; the stellar interior is opaque, or optically thick.

Nuclear energy is released in the interiors of stars and flows outward through their interiors and then their atmospheres. As radiation passes through the matter in both regions, photons are emitted, absorbed, and re-emitted; the directions and/or the wavelengths of these re-emitted photons may be altered from their original states. It is this process of **radiative transfer** that we will study first.

Later chapters will deal with the structure of the atom, excitation, ionization, the interpretation of stellar spectra, line profiles, and sources of opacity. Once these basic ideas are in hand, they will be applied to the tasks of determining how stars are structured, how they function, and how they form, evolve, and die.

CHAPTER 2: Radiative Transfer

Radiation Terminology

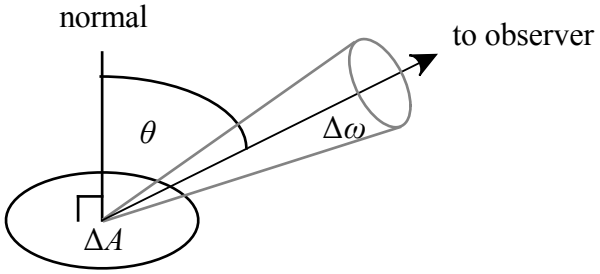
The basic problem of radiative transfer involves the passage of radiation through matter. As it does so, some of the radiation is absorbed, and some radiation is emitted. The rate at which radiation is absorbed by matter will depend on the rate at which radiation is incident on the matter – more radiation results in more absorption; the rate at which radiation is emitted will depend only slightly on the incident radiation. The question is how the radiation changes as it passes through matter.

Before we can find an answer, we have to decide what property of the radiation we want to monitor. There are several possibilities, and they all begin with a quantity called the **specific intensity**, $\boldsymbol{I_\nu}$.

Specific Intensity

Consider rays of light of intensity I_ν, passing through a surface element of area ΔA, towards the observer at an angle θ to the normal (of the surface element), into **solid angle**[*] $\Delta\omega$, as shown in Figure 2.1.

Figure 2.1: A ray of light passing through a surface element ΔA into a solid angle Δω at an angle θ to the normal



Let ΔE_ν equal the amount of energy in photons with frequency in the range ν to $\nu + \Delta\nu$, passing through the area ΔA into solid angle $\Delta\omega$ in a time Δt. Then the **specific intensity** $\boldsymbol{I_\nu}$ is defined by the following:

[*] See Appendix: Solid Angle.

Eq. 2.1 $$I_\nu \equiv \lim_{\Delta\to 0} \frac{\Delta E_\nu}{\cos\theta\, \Delta A\, \Delta t\, \Delta\nu\, \Delta\omega}$$ (Note: $\Delta A \cos\theta$ is the *projected area*.)

or $$I_\nu \equiv \frac{dE_\nu}{\cos\theta\, dA\, dt\, d\nu\, d\omega}$$

The limit here is a practical limit rather than a strictly mathematical limit. This means we will shrink our Δ's down such that we are considering regions smaller than our resolving power, but not so small as to cause problems with the **uncertainty principle** or other quantum mechanical factors.

The dimensions of I_ν can be obtained from the definition: in cgs units – the standard for astronomy – E is in ergs, A is in square centimeters, t is in seconds, ν is in hertz, and ω is in steradians, making the dimensions of I_ν $\frac{\text{ergs}}{\text{cm}^2\text{-s-Hz-st}}$. Note: The reader might observe that multiplication of seconds by hertz would cancel both units, thus simplifying the expression; however, the reader should strongly resist the temptation to do this.

In general, I_ν is a function of position in space, direction (of the ray), and frequency; it is a measure of the brightness of a particular ray of a particular frequency at a particular point along the ray. I_ν can be defined for a point on a radiating surface or for a point in space (with a specified direction).

Note that once we have taken the limit as the solid angle ($\Delta\omega$) goes to zero, we have a *ray*, which is not divergent. This means that the energy is not being spread out over a range of angles and diluted by distance, and hence, I_ν is independent of distance from the source – I_ν is constant along a ray. The intensity of a ray of sunlight is thus constant: I_ν for a point on the Sun would be the same as measured from any planet. (Note: If this seems counter-intuitive, it is because there is *another* property of radiation that *does* diminish with distance; stay tuned.)

Astronomers use telescopes to make stars appear brighter, but they do this not by increasing the intensity of a ray, but rather by collecting *more* rays from the source. Note also that intensity can only be measured for sources of finite angular size, such as the Sun, the Moon, or the planets, but it cannot be measured for point sources, such as stars. Why not?

I_ν depends on the angle θ between the ray to the observer and the normal to the surface at which the intensity is measured; different rays passing through the same point but going in different directions may have different intensities. But sometimes we do not care which direction the rays are moving and only want to know the average value of the intensity at a point. In this case, we can average I_ν over all angles (over a sphere – over 4π) to get the mean intensity.

Mean Intensity

To find the **mean intensity**, $\boldsymbol{J_\nu}$, we perform a weighted average. We find the specific intensity, I_ν, for each differential solid angle ($d\omega$), multiply I_ν by $d\omega$, and add these products up for the whole sphere. Then we must divide this sum by the sum of the weights – the sum of the

differential solid angles – to obtain an average value for the intensity. Mathematically, this becomes the following:

Eq. 2.2 $$J_\nu \equiv \frac{\int_{4\pi} I_\nu d\omega}{\int_{4\pi} d\omega} = \frac{1}{4\pi}\int_{4\pi} I_\nu d\omega$$

Note here that the differential solid angle can be written as $d\omega = \sin\theta\, d\theta\, d\phi$, and the integral in the denominator then can be easily solved:

Eq. 2.3 $$\int_{4\pi} d\omega = \int_{4\pi} \sin\theta\, d\theta\, d\phi = \int_0^{2\pi}\int_0^{\pi} \sin\theta\, d\theta\, d\phi = 2\pi\int_0^{\pi}\sin\theta\, d\theta = 2\pi[\cos\theta]_\pi^0 = 4\pi$$

J_ν is an intensity, as is I_ν, and therefore both quantities have the same units, described above for I_ν. While I_ν can vary with the direction of the ray, J_ν is an average value over all angles, and therefore it does not vary with direction. If I_ν should happen to be **isotropic** – meaning 'direction independent' – then I_ν can be moved outside the integral to obtain the following:

Eq. 2.4 $$J_\nu = \frac{1}{4\pi}\int_{4\pi} I_\nu d\omega = \frac{1}{4\pi} I_\nu \int_{4\pi} d\omega = I_\nu$$

Radiative Flux

Sometimes the question of interest is a bit different: How much energy in an arbitrary radiation field will pass through an area ΔA in a time Δt in a frequency bandwidth $\Delta\nu$? (Note: Here we have no mention of angles.) The answer to this question will be the **radiative flux** (or simply, the **flux**) $\boldsymbol{F_\nu}$, defined as follows:

Eq. 2.5 $$\Delta F_\nu = \frac{\Delta E_\nu}{\Delta A\, \Delta t\, \Delta\nu} = \frac{\Delta E_\nu}{\cos\theta\, \Delta A\, \Delta t\, \Delta\nu\, \Delta\omega}\cos\theta\, \Delta\omega$$

Then taking the limit, we find

Eq. 2.6 $$dF_\nu = \lim_{\Delta\to 0}\Delta F_\nu = \lim_{\Delta\to 0}\frac{\Delta E_\nu}{\cos\theta\, \Delta A\, \Delta t\, \Delta\nu\, \Delta\omega}\cos\theta\, \Delta\omega = I_\nu \cos\theta\, d\omega$$

From this, we find the **flux** to be

Eq. 2.7 $$F_\nu = \int_{4\pi} I_\nu \cos\theta\, d\omega$$

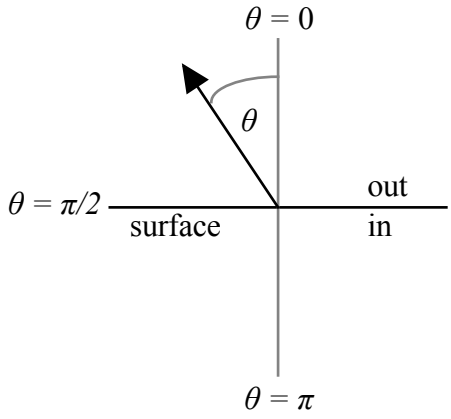
The units of this flux are intensity units multiplied by solid angle units, or $\frac{\text{ergs}}{\text{cm}^2\text{-s-Hz}}$.

Now if I_ν is isotropic – a fairly common simplifying assumption – then equal amounts of energy will pass through the area in each direction and the flux (or **net flux**) will be zero, as integration confirms:

Eq. 2.8 $$F_\nu = I_\nu\int_{4\pi}\cos\theta\, d\omega = I_\nu\int_0^{2\pi}\int_0^{\pi}\cos\theta\sin\theta\, d\theta\, d\phi = 2\pi I_\nu\int_0^{\pi}\sin\theta\cos\theta\, d\theta$$
$$= \pi I_\nu[\sin^2\theta]_0^{\pi} = 0$$

In performing these integrations, the convention is to label the outward normal as $\theta = 0$, as shown in Figure 2.2.

Figure 2.2: Angular coordinate convention at the stellar surface



Sometimes only the **outward flux** (or emittance) is desired. This can be found the same as in Equation 2.7 but integrating over only the outward hemisphere, which can be accomplished by changing the limits of integration.

Eq. 2.9 $$F_\nu^+ = \int_{2\pi} I_\nu \cos\theta \, d\omega$$

And if I_ν is isotropic, the integral can proceed easily:

Eq. 2.10 $$F_\nu^+ = I_\nu \int_{2\pi} \cos\theta \, d\omega = I_\nu \int_0^{2\pi} \int_0^{\pi/2} \cos\theta \sin\theta \, d\theta \, d\phi = 2\pi I_\nu \int_0^{\pi/2} \sin\theta \cos\theta \, d\theta$$
$$= \pi I_\nu [\sin^2\theta]_0^{\pi/2} = \pi I_\nu$$

There are some who would prefer to eliminate the factor of π that appears in the result for Equation 2.10 ($F_\nu^+ = \pi I_\nu$), and this has been done by defining a quantity ($\mathcal{F}_\nu \equiv F_\nu / \pi$) most generally known as the **astrophysical flux** (see Gray (1976), Novotny (1973), Bohm-Vitense (1989)), but also as the radiative flux (see Collins (1989)). Bowers and Deeming (1984) have the same two quantities but reverse the notation: their flux is $\mathcal{F}_\nu$ while their astrophysical flux is F_ν, making their linking equation $\mathcal{F}_\nu = \pi F_\nu$. In this book, we will not use the astrophysical flux.

There are some occasions where it is useful to define a flux in the same manner as we defined the mean intensity – as a weighted average of a quantity. In this case, the quantity to average is the product $I_\nu \cos\theta$. The result is $\boldsymbol{H_\nu}$ – the **Harvard flux** or **Eddington's flux**.

Eq. 2.11 $$H_\nu \equiv \frac{\int_{4\pi} I_\nu \cos\theta \, d\omega}{\int_{4\pi} d\omega} = \frac{1}{4\pi} \int_{4\pi} I_\nu \cos\theta \, d\omega = \frac{F_\nu}{4\pi}$$

Radiation Pressure

Photons exert pressure – **radiation pressure**, designated $\boldsymbol{P_\nu}$. Radiation pressure is responsible for the dust tails of comets, as it pushes the tiny dust particles released by the comet

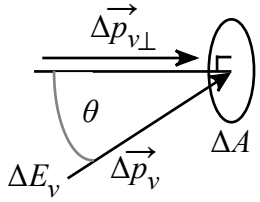
into higher orbits about the Sun. It also makes an important contribution to the pressure at very high temperatures in stars.

Radiation pressure can be calculated for a particular intensity as follows. In general, pressure is equal to force per unit area, with only the component of the force that is perpendicular to the area being counted. But because force is the time derivative of momentum, pressure is then equal to the **momentum flux**.

Eq. 2.12
$$\text{pressure} = \frac{\perp \text{ force}}{\text{area}} = \frac{\frac{\text{d}}{\text{dt}} \perp \text{momentum}}{\text{area}} = \text{momentum flux} \left[\frac{\text{momentum}}{\text{cm}^2\text{-s}} \right]$$

The momentum p_ν of a group of photons with energy E_ν is $p_\nu = E_\nu$ /c. Addition of photons with energy ΔE_ν to the group changes the momentum by $\Delta p_\nu = \Delta E_\nu$ /c.

Figure 2.3: Photon momentum transfer to a surface



Consider photons with energy ΔE_ν that are incident on a surface area ΔA at an angle to the normal θ, as shown in Figure 2.3. These photons will transfer momentum to the surface, but only the perpendicular component of the momentum – given in Equation 2.13 – will contribute to the radiation pressure.

Eq. 2.13
$$\Delta p_{\nu\perp} = \frac{\Delta E_\nu}{c} \cos\theta \qquad (c \text{ is the speed of light}^*)$$

The perpendicular momentum flux (the radiation pressure) of these photons is then as follows:

Eq. 2.14
$$\Delta P_\nu = \frac{\Delta p_{\nu\perp}}{\Delta A\, \Delta t\, \Delta\nu} = \frac{\Delta E_\nu \cos\theta}{c\, \Delta A\, \Delta t\, \Delta\nu} \times \frac{\cos\theta\, \Delta\omega}{\cos\theta\, \Delta\omega}$$

Multiplying both numerator and denominator by cos θ $\Delta\omega$ and taking the limit yields the following:

Eq. 2.15
$$dP_\nu = \lim_{\Delta\to 0} \Delta P_\nu = \lim_{\Delta\to 0} \frac{\Delta E_\nu}{\cos\theta\, \Delta A\, \Delta t\, \Delta\nu\, \Delta\omega} \frac{\cos^2\theta\, \Delta\omega}{c}$$
$$= \frac{I_\nu}{c} \cos^2\theta\, d\omega$$

Finally, integrating this expression for dP_ν over all directions gives the radiation pressure P_ν.

---

* See Appendix: Constants

Eq. 2.16 $$P_\nu = \int_{4\pi} dP_\nu = \frac{1}{c}\int_{4\pi} I_\nu \cos^2\theta \, d\omega \quad \left[\frac{\text{dynes}}{\text{cm}^2\text{-Hz}}\right]$$

Note that the radiation pressure integral contains two factors of cos θ : one is because we want the projected area of the surface element, and the other is because only the perpendicular component of the momentum contributes to the pressure.

Equation 2.16 gives the monochromatic radiation pressure (good for photons in the range ν to $\nu + d\nu$). A related integral is defined by again taking a weighted average:

Eq. 2.17 $$K_\nu \equiv \frac{\int_{4\pi} I_\nu \cos^2\theta \, d\omega}{\int_{4\pi} d\omega} = \frac{1}{4\pi}\int_{4\pi} I_\nu \cos^2\theta \, d\omega = \frac{cP_\nu}{4\pi} \quad \text{and} \quad P_\nu = \frac{4\pi}{c}K_\nu$$

Now if I_ν is isotropic, P_ν can be calculated as follows:

Eq. 2.18 $$P_\nu = \frac{I_\nu}{c}\int_{4\pi} \cos^2\theta \, d\omega = \frac{2\pi}{c} I_\nu \left(\tfrac{1}{3}\cos^3\theta\right|_\pi^0 = \frac{4\pi}{3c} I_\nu$$

Note the 'moments of I_ν': these are convenient integrals to use because they share a similar form.

Eq. 2.19a $$J_\nu = \frac{1}{4\pi}\int I_\nu \, d\omega$$

Eq. 2.19b $$H_\nu = \frac{1}{4\pi}\int I_\nu \cos\theta \, d\omega$$

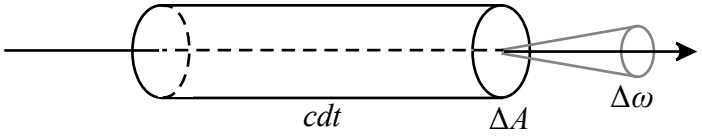
Eq. 2.19c $$K_\nu = \frac{1}{4\pi}\int I_\nu \cos^2\theta \, d\omega$$

Radiative Energy Density

A radiation field contains photons traveling in various directions through space. As each photon carries energy, there will be some radiative energy per unit volume of space; this quantity is the **radiative energy density u_ν** . Its value can be determined by carefully selecting a volume of space and then counting the number of photons within it.

Let us begin by selecting a ray of light and then construct around it a cylindrical volume dV of length cdt and cross-sectional area dA, as shown in Figure 2.4.

Figure 2.4: Volume of space used to determine radiative energy density



Now consider all of the photons within dV (= $c\,dt\,dA$) that have a frequency in the range ν to $\nu + d\nu$ and are directed into a solid angle $d\omega$. If the energy density is $u_\nu(\omega)$, then the energy contained in dV is given by the following:

Eq. 2.20 $dE_\nu = u_\nu(\omega)\, dV\, d\nu\, d\omega = u_\nu(\omega)\, dA\, cdt\, d\nu\, d\omega$

Because of the manner in which this volume was constructed, the energy described above must all flow out of the volume through area dA into solid angle $d\omega$ in time dt. This energy is as follows:

Eq. 2.21 $dE_\nu = I_\nu\, dA\, dt\, d\nu\, d\omega$ (Note: $\cos\theta = 1$)

Equating these two energies yields a relation between energy density and specific intensity:

Eq. 2.22 $$u_\nu(\omega) = \frac{I_\nu}{c}$$

Then the energy density can be found by integration:

Eq. 2.23 $$u_\nu = \int_{4\pi} u_\nu(\omega)\, d\omega = \frac{1}{c}\int_{4\pi} I_\nu\, d\omega = \frac{4\pi}{c} J_\nu \quad \left[\frac{\text{ergs}}{\text{cm}^3\text{-Hz}}\right]$$

If I_ν is isotropic, then $J_\nu = I_\nu$, and $u_\nu = {}^{4\pi}/_{c}\, I_\nu$. But because $P_\nu = {}^{4\pi}/_{3c}\, I_\nu$ in the isotropic case, then $P_\nu = {}^{1}/_{3}\, u_\nu$. In an isotropic radiation field, radiation pressure is one third of the energy density. Note: The units of pressure (dynes/cm^2) and energy density (ergs/cm^3) are the same.

Thus, we have several ways of characterizing the radiation field, but all are ultimately dependent on the specific intensity, which is a function of position, direction, and frequency. Next we must turn our attention to the interactions between radiation and matter.

Radiation and Matter

As noted above, the intensity of a ray of light does not change as it passes through the vacuum of space. But if light encounters matter along the way, the intensity may be altered by interactions between the radiation field and the matter. (The details of such interactions will be explored in a later chapter.) In general, the intensity may be decreased (by photons being either absorbed or scattered out of the ray) or increased (by photons being either emitted or scattered into the ray). The rates at which these changes occur are defined in terms of coefficients specific to each process.

Emission Coefficient

Atoms may emit light spontaneously, without any particular provocation. We can account for this process by defining the **spontaneous emission coefficient j_ν** as the energy emitted per unit volume per unit time per frequency interval per solid angle.

Eq. 2.24 $$dE_\nu = j_\nu\, dV\, dt\, d\nu\, d\omega \quad \Rightarrow \quad j_\nu = \frac{dE_\nu}{dV\, dt\, d\nu\, d\omega}$$

The units of this *volume* emission coefficient j_ν are then $\left[\frac{\text{ergs}}{\text{cm}^3\text{-s-Hz-st}}\right]$. The volume emission coefficient is used by Rybicki & Lightman (1979) and by Bohm-Vitense (1989), although the latter designates it as ε_ν rather than j_ν.

Other authors (Gray (1976), Collins (1989)) define the emission coefficient in a different manner, as the amount of energy emitted *per gram* per unit time per frequency interval per solid angle.

Eq. 2.25 $$dE_\nu = j_\nu \, dt \, d\nu \, d\omega \quad \Rightarrow j_\nu = \frac{dE_\nu}{dt \, d\nu \, d\omega}$$

(Note that although this coefficient is *per gram*, there is no *dm* factor in the expression.)

The units of this *mass* emission coefficient $\left[\frac{\text{ergs}}{\text{g-s-Hz-st}} \right]$ are slightly different from those of the *volume* emission coefficient. However, the same symbol (j_ν) is used for both coefficients. This may seem potentially confusing, but in reality, it is not that hard to tell which coefficient is meant. And most authors will usually focus on only one of these; in this book, we will normally use the mass emission coefficient.

In either case, the *isotropic* emission coefficient is then given by $\int_{4\pi} j_\nu \, d\omega = 4\pi \, j_\nu$ (Novotny, 1973), but there is no special symbol for this term. Spontaneous emission is inherently isotropic; atoms have no preferred direction for emission of photons.

The emission coefficient appears as follows: Consider a beam of light of intensity I_ν passing through a length (dx) of matter of density ρ (g/cm^3). (The signs are set such that x increases along the ray.) The change in intensity (dI_ν) due to emission is then given by the following:

Eq. 2.26a $$dI_\nu = j_\nu \, \rho \, dx$$ or

Eq. 2.26b $$dI_\nu = j_\nu \, dx$$ for the volume coefficient

Note that spontaneous emission is independent of the value of I_ν , which may even be zero; radiation need not be present in order for spontaneous emission to occur.

Absorption Coefficient

Absorption is a different matter, as without photons to absorb, there can be no absorption. Thus, absorption *requires* the presence of a radiation field, and the *rate* of absorption will depend on the value of the intensity of the radiation: with more photons available, there will be more absorption.

Experimentally, this is found to be the case. If conditions are such that negligible emission occurs, the intensity of a beam of radiation is found to diminish exponentially as it passes through matter, according to the following:

Eq. 2.27 $$\frac{dI_\nu}{dx} \approx -I_\nu$$

Here the negative sign signifies that the intensity decreases as x increases – that is, as the ray travels deeper into the matter. This relation can be used to define an absorption coefficient κ_ν , as follows:

Eq. 2.28 $$dI_\nu = -\kappa_\nu \, \rho \, I_\nu \, dx$$

The units of κ_ν can be determined from this equation; the intensity units on either side cancel each other, leaving the units of κ_ν equal to those of dx/ρ: cm/(g/cm^3) or cm^2/g. These units imply a simple, intuitive meaning for κ_ν ; if we consider that each atom provides a target for the photons, then each gram of atoms would provide a collective target area equal to the value of κ_ν. The greater this target area, the more likely absorption will occur. Thus, κ_ν is a *mass* absorption coefficient, giving the absorption target area per unit mass.

We can also define a *volume* absorption coefficient α_ν, which relates to κ_ν through the density ρ:

Eq. 2.29 $$\alpha_\nu = \kappa_\nu \rho$$

And it relates to the intensity in a manner similar to Equation 2.28:

Eq. 2.30 $$dI_\nu = -\alpha_\nu I_\nu dx$$

The units of α_ν can be arrived at in several ways. If κ_ν is the target area per unit mass (cm^2/g), then α_ν is the target area per unit volume: cm^2/cm^3, which of course reduces to cm^{-1}. This same result can be obtained by multiplying the units of κ_ν and ρ together, or by noting from Equation 2.30 that α_ν must have the inverse units of x.

Radiative transfer may be done on a per gram basis, using j_ν and κ_ν, or on a per volume basis, using j_ν and α_ν. As noted, we will normally utilize the former, but will stay alert for occasions better suited to the latter.

Cross Section

So far we have discussed the absorption coefficient as a target area per unit mass or as a target area per unit volume, but it can also be considered as a target area per particle; that is, how large an absorption target does each particle in the gas present? When used in this context, the absorption coefficient is usually termed a **cross section**. (Cross sections may be employed for other uses as well, such as in nuclear reactions.)

The most common symbol for the cross section is σ_ν, and it has units of cm^2/particle or just cm^2. To determine the total cross section for a gas, we would then need to know the **number density** of particles – ***n*** (particles/cm^3). Multiplying the number density (particles/cm^3) by the cross section (cm^2/particle) yields units of cm^{-1}, the same as α_ν. Thus, the cross section relates to the two absorption coefficients as follows:

Eq. 2.31 $$\alpha_\nu = \kappa_\nu \rho = \sigma_\nu n$$

Mean Free Path

A related concept in this discussion of absorption is the question of how far a photon travels through matter, on average, before it is absorbed – a distance called the **mean free path, ℓ_ν** . This quantity is inversely related to the cross section and the absorption coefficients: as the cross

section (or absorption coefficient) increases, the mean free path decreases, as the following shows.

Eq. 2.32 $$\ell_\nu = 1/\kappa_\nu \rho \;\; = 1/\sigma_\nu n \;\; = 1/\alpha_\nu$$

Nomenclature

There are a variety of different terms that are used in connection with absorption coefficients, including cross section and **opacity**. The opacity is the probability of a photon being absorbed (or scattered) by matter. This probability can be expressed in several different ways, producing the different coefficients that we have already learned.

α_ν is the probability of absorption per unit path length, in cm^{-1}.

α_ν is also the cross section per unit volume, in cm^2/cm^3.

κ_ν is the cross section per unit mass, in cm^2/g.

σ_ν is the cross section per particle, in cm^2/particle.

Table 2.1: Opacity nomenclature

Author	• Volume Opacity • Total Extinction Coefficient • Absorption Coefficient	• Specific Opacity • (Mass) Absorption Coefficient • Opacity Coefficient • Mass Scattering Coefficient	• Opacity per Particle • Cross section • Atomic Absorption Coefficient
Pierce (2013)	α_ν	κ_ν	σ_ν
Bowers & Deeming (1984)	k_ν	κ_ν	σ_ν
Gray (1976)	–	κ_ν	κ, α
Novotny (1973)	–	κ_ν , σ	a
Rybicki & Lightman (1979)	α_ν	κ_ν	σ_ν
Bohm-Vitense (1989)	κ_λ	–	σ
Collins (1989)	α	κ_ν , σ_ν	α_ν
Carroll & Ostlie (1996)	–	κ_λ	σ

Not every author uses each of these coefficients, nor do they all agree on which symbols to use for them. Table 2.1 illustrates the variation that can be expected in the literature. (Note the use of the subscript λ, in place of ν in two instances; the quantities being discussed may also be written in terms of wavelength, rather than frequency, and some authors prefer this variation.)

This author will attempt to employ consistent usage throughout this book.

The Transfer Equation for Normal Rays

Now, armed with appropriate coefficients for absorption and emission, we are ready to allow for both processes to occur as light passes through matter along a normal ray – one perpendicular to the surface. We may begin by combining Equations 2.26 and 2.28 to obtain the following:

Eq. 2.33 $$dI_\nu = -\kappa_\nu \rho \, I_\nu \, dx + j_\nu \rho \, dx$$

The equivalent 'per volume' equation is also presented here for comparison:

Eq. 2.34 $$dI_\nu = -\alpha_\nu I_\nu \, dx + j_\nu \, dx$$

The **radiative transfer equation** then follows from Equation 2.33:

Eq. 2.35 $$\frac{dI_\nu}{dx} = -\kappa_\nu \rho I_\nu + j_\nu \rho$$

There are many forms of the radiative transfer equation, with some better suited to solving particular problems and others containing more appropriate variables. We will introduce some new variables now by dividing through both sides of the equation by $\kappa_\nu \rho$ to obtain the following:

Eq. 2.36 $$\frac{dI_\nu}{\kappa_\nu \rho \, dx} = -I_\nu + \frac{j_\nu}{\kappa_\nu}$$

Now define the **source function S_ν** as the ratio of emission coefficient to absorption coefficient:

Eq. 2.37 $$S_\nu \equiv \frac{j_\nu}{\kappa_\nu}$$

And define the **differential optical depth $d\tau_\nu$** :

Eq. 2.38 $$d\tau_\nu \equiv \kappa_\nu \rho \, dx$$

Inserting these variables into the transfer equation gives the following simplified version:

Eq. 2.39 $$\frac{dI_\nu}{d\tau_\nu} = -I_\nu + S_\nu$$

It should be clear from this equation that the source function has the same units as intensity, and that optical depth is dimensionless.

Optical Depth

The **optical depth τ_ν** gives the overall effectiveness of a given amount of matter at absorbing photons. It is a measure of the amount of matter the ray must path through, combined with the opacity, which specifies the matter's effectiveness at absorbing photons. Both factors are needed: a small thickness of matter with a high opacity can have the same optical depth as a huge cloud of matter with a low opacity.

The differential equation 2.38 can be integrated to give an expression for τ_ν :

Eq. 2.40 $$\tau_\nu(x) = \int_0^x \kappa_\nu \rho \, dx'$$ (x' is a dummy integration variable.)

Here the optical depth and the path length are both zero where the ray enters the matter, and they increase together along the path. That is, $\tau_\nu(0) = 0$, and because κ_ν and ρ are both positive quantities, τ_ν and x will increase at the same time. If κ_ν and ρ are both constant throughout the medium, then the expression for the optical depth is further simplified:

Eq. 2.41 $\tau_\nu(x) = \kappa_\nu \rho\, x$

If the value of τ_ν remains fairly low when the ray has passed completely through the matter ($\tau_\nu << 1$), we say the medium is *optically thin*, or transparent. This means that the average photon can travel through the medium without being absorbed. On the other hand, if the value of τ_ν becomes relatively high as the ray passes through the matter ($\tau_\nu >> 1$), we say the medium is *optically thick*, or opaque. This means that the average photon will be absorbed – perhaps many times – as it attempts to travel through the medium.

Optical depth provides a way to characterize the optical properties of matter, without necessarily specifying all of the details of the physical properties. Matter that is opaque may have a large absorption coefficient, a high density, and/or a long path length for the photons to travel.

The Source Function

The **source function S_ν** is a property of the matter in its particular thermodynamic state. For example, two gases with the same composition but at different temperatures and/or densities would have different source functions. Because the source function is the ratio of the emission coefficient to the absorption coefficient, it includes the effects of emission and absorption – and *scattering*, too.

A scattered photon has the same energy as the incoming photon, but its direction may be changed by the encounter. For our purposes, scattering can be treated as absorption followed immediately by emission of a photon of the same energy in a random direction; thus, we need only expand our meanings of j_ν and κ_ν to include scattering, and we will be able to use our existing transfer equation, without adding a third term.

In κ_ν we will include photons that have been absorbed (with their energy *thermalized* – turned into thermal energy of the particles of absorbing matter) and also photons that have been scattered *out of* the beam solid angle by the matter.

In j_ν we will include photons that have been emitted (with their energy coming from the thermal energy of the particles of absorbing matter) and also photons that have been scattered *into* the beam solid angle.

Let us now illustrate some special, simple cases of the source function, which represent opposite extremes of the contribution of scattering: **pure scattering** and **pure absorption**.

In the case of pure scattering, we will suppose that all of the emitted energy represented by the emission coefficient j_ν is due to photons being scattered into the beam by the matter. It is not necessary to assume that *every* scattered photon is scattered into the beam – only that scattering is the only source of new photons in the beam.

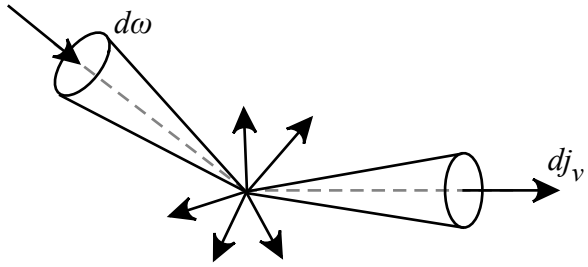
In keeping with our treatment of scattering above, we may say that photons are being absorbed and then immediately re-emitted isotropically at the same frequency, with some fraction of these being in the beam of interest. What is the appropriate source function?

Figure 2.5 illustrates photons encountering a particle of matter from some random solid angle $d\omega$ and being scattered into the beam, where they are counted towards the emission as dj_ν. The

portion of the incoming intensity I_ν that is scattered is determined by the absorption coefficient κ_ν, allowing us to write the following expression for the contribution to the beam emission dj_ν from the solid angle $d\omega$:

Eq. 2.42 $$dj_\nu = \kappa_\nu I_\nu d\omega/4\pi$$

Figure 2.5: Scattering of photons into the beam from solid angle dω



Here the factor $\kappa_\nu I_\nu$ represents the rate at which photons from all directions are being absorbed by the matter, and the ratio $d\omega/4\pi$ is the fraction of those absorbed photons that came from the solid angle $d\omega$. To obtain the total beam emission, we must integrate the previous equation over all incoming solid angles:

Eq. 2.43 $$j_\nu = \frac{1}{4\pi}\int_{4\pi} \kappa_\nu I_\nu d\omega$$

But because κ_ν is generally independent of direction (atoms have no preferred direction from which they absorb photons), we may remove this coefficient from the integral, which leads to an easy solution:

Eq. 2.44 $$j_\nu = \frac{\kappa_\nu}{4\pi}\int_{4\pi} I_\nu d\omega = \kappa_\nu J_\nu$$

And the source function is then very simple:

Eq. 2.45 $$S_\nu \equiv \frac{j_\nu}{\kappa_\nu} = J_\nu$$

That is, in the case of pure scattering, the source function is equal to the mean intensity. This means that the photons emerging from the matter will have the same characteristics as the photons that entered the matter, except for some redirection. In particular, the *spectral* distribution of the photons is not modified by their passage through the matter.

The other special case of note is that of **pure absorption**, in which it is assumed that no scattering takes place. Incoming photons are absorbed by the matter and their energy is thermalized; new photons are then created from this pool of energy and emitted; their spectral distribution will be governed completely by the physical state of the matter from which they emerge.

Discussion of this process will normally be limited to a particular state of matter called **local thermodynamic equilibrium (LTE)**. This means that within a local region of apace, matter has reached a state of equilibrium in which no *net* transitions are occurring; the distribution of

particles in different energy states is given by standard distribution functions (to be discussed later) that are characterized by the temperature.

For example, in the case of pure absorption, if the matter is opaque, then the radiation it emits will have the spectrum of a **blackbody**, and the source function will be the photon distribution function associated with blackbody radiation, called the **Planck function**, $\boldsymbol{B_\nu(T)}$.

Eq. 2.46 $$B_\nu(T) = \frac{2h\nu^3}{c^2}\frac{1}{e^{h\nu/kT}-1} \qquad \text{with units} \left[\frac{\text{ergs}}{\text{cm}^2\text{-s-Hz-st}}\right]$$

The Planck function is often written in terms of wavelength, rather than frequency, in which case it is known as $\boldsymbol{B_\lambda(T)}$.

Eq. 2.47 $$B_\lambda(T) = \frac{2hc^2}{\lambda^5}\frac{1}{e^{hc/\lambda kT}-1} \qquad \text{with units} \left[\frac{\text{ergs}}{\text{cm}^2\text{-s-cm-st}}\right]$$

(Once again, the reader should strongly resist the temptation to combine these units.)

It should be noted that although $B_\nu(T)$ and $B_\lambda(T)$ have similar names and similar forms, they really are *different* functions, with *different* units, and thus are not interchangeable. They are, of course, related; when we integrate $B_\nu(T)$ over all frequencies, we should get the same result as when we integrate $B_\lambda(T)$ over all wavelengths – a result we call $\boldsymbol{B(T)}$, the **integrated Planck function**.

Eq. 2.48 $$\int_0^\infty B_\nu(T)\,d\nu = \int_0^\infty B_\lambda(T)\,d\lambda = B(T)$$

The differential form of this equation

Eq. 2.49 $$dB(T) = B_\nu(T)d\nu = B_\lambda(T)d\lambda$$

leads to the transformation expression

Eq. 2.50 $$B_\lambda(T) = B_\nu(T)\,d\nu/d\lambda \qquad \text{with } \nu = c/\lambda \quad \text{and} \quad d\nu/d\lambda = -c/\lambda^2$$

The Planck function has a standard form, shown in Figure 2.6. $B_\nu(T)$ plotted vs. frequency exhibits a similar form. In either case, the curve starts at zero, rises to a maximum and then falls off, becoming asymptotic to the horizontal axis. Increasing the temperature will move each point higher and shift the peak to shorter wavelengths. The wavelength where this peak occurs is found by setting the derivative of the Planck function to zero. The result is known as **Wien's displacement law**:

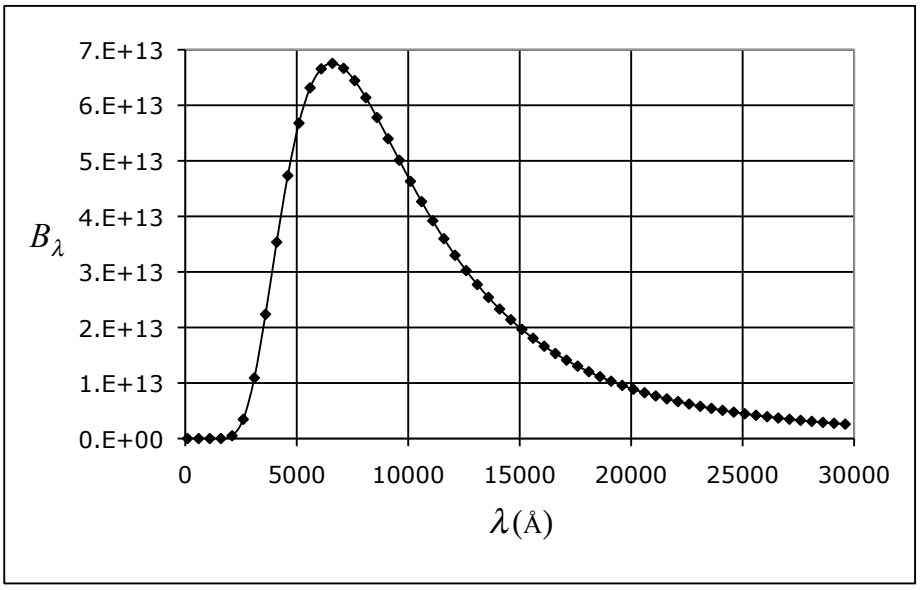
Eq. 2.51 $$\lambda_{\max} = \frac{\text{constant}}{T} \qquad \text{where the constant is about 0.2898 cm-K}$$

As noted above, integrating either Planck function gives $B(T)$, which is the area under the curve. The value of the integrated Planck function is as follows:

Eq. 2.52 $$B(T) = \frac{\sigma T^4}{\pi} \qquad \text{with units} \left[\frac{\text{ergs}}{\text{cm}^2\text{-s-st}}\right]$$

(Here σ is the Stefan-Boltzmann constant*, not a cross section.)

Figure 2.6: The Planck function, $B_\lambda(T)$ for T = 4400 K



We could use these two extreme cases of pure scattering and pure absorption to arrive at a suitable source function by defining different opacities for scattering (κ_ν^S) and absorption (κ_ν^A), with a total opacity (κ_ν) equal to their sum. Then the emission coefficient, which is generally the product of the opacity and the source function, can be written as a combination of terms, as follows:

Eq. 2.53 $$j_\nu = \kappa_\nu^S J_\nu + \kappa_\nu^A B_\nu(T)$$

And the source function, which is j_ν / κ_ν, can then be calculated:

Eq. 2.54 $$S_\nu = \frac{\kappa_\nu^S}{\kappa_\nu} J_\nu + \frac{\kappa_\nu^A}{\kappa_\nu} B_\nu(T)$$

For now however, we will keep things simple and just use κ_ν for our opacity.

Solutions of the Transfer Equation

It is now time to find a solution to the transfer equation. Because the transfer equation is a differential equation with I_ν as the dependent variable, the solution we seek will be an equation giving I_ν as a function of the independent variable, usually x or τ_ν. That is, we want to know how the intensity varies as it passes through matter.

To begin, we will pick a suitable form of the transfer equation, perhaps Equation 2.35:

Eq. 2.35 $$\frac{dI_\nu}{dx} = -\kappa_\nu \rho I_\nu + j_\nu \rho$$

* See Appendix: Constants

The independent variable is x, the path length, which increases along the ray.

First we will investigate the special case in which *no absorption* occurs. This can be accomplished by setting κ_ν equal to zero. Then the transfer equation is much simpler:

Eq. 2.55 $$\frac{dI_\nu}{dx} = j_\nu \rho$$

The integral form of this equation is as follows:

Eq. 2.56 $$\int dI_\nu = \int j_\nu \rho \, dx$$

And the solution can be written in this form:

Eq. 2.57 $$I_\nu(x) = I_\nu(0) + \int_0^x j_\nu \rho \, dx' \qquad \textit{special case solution: no absorption}$$

Because the integral term in this solution is positive for all points in the matter, we can say that the intensity must *increase* along the ray. We cannot proceed any further without knowing how the emission coefficient and/or the density vary along the path.

The next special case to consider is that with *no emission*, which we can obtain by setting j_ν equal to zero. The transfer equation again simplifies:

Eq. 2.58 $$\frac{dI_\nu}{dx} = -\kappa_\nu \rho I_\nu$$

This differential equation can be solved by first separating variables:

Eq. 2.59 $$\frac{dI_\nu}{I_\nu} = -\kappa_\nu \rho \, dx$$

This can be easily integrated:

Eq. 2.60 $$\ln I_\nu(x) - \ln I_\nu(0) = \int_0^x \kappa_\nu \rho \, dx'$$

Or, in a more common form:

Eq. 2.61 $$I_\nu(x) = I_\nu(0) \, e^{-\int_0^x \kappa_\nu \rho \, dx'}$$

Or, in a simpler form:

Eq. 2.62 $$I_\nu(\tau_\nu) = I_\nu(0) \, e^{-\tau_\nu} \qquad \textit{special case solution: no emission}$$

From this, it should be clear that if no emission is occurring, the intensity will *decrease* exponentially along the ray, with the optical depth being the independent variable. Of course, it would be best to know how the intensity varies if *both* emission and absorption are allowed to proceed, as that situation is most apt to present itself in real life. Unfortunately, the solution will not be obtained as simply as by combining the two special case solutions. Instead, we must return to the transfer equation, this time beginning with Equation 2.39:

Eq. 2.39 $$\frac{dI_\nu}{d\tau_\nu} = -I_\nu + S_\nu$$

For this equation, both the intensity and the source function are functions of the optical depth, which increases along the ray. Let us begin our solution of this differential equation by rewriting it:

Eq. 2.63 $$\frac{dI_\nu}{d\tau_\nu} + I_\nu = S_\nu$$

We will now employ one of the standard tricks used to solve differential equations: multiplying both sides of the equation by a common factor – in this case, by e^{τ_ν}.

Eq. 2.64 $$e^{\tau_\nu}\frac{dI_\nu}{d\tau_\nu} + I_\nu e^{\tau_\nu} = S_\nu e^{\tau_\nu}$$

At this point, the astute reader will note that we now have an *exact differential*; that is, the left side of the equation is the derivative of some function, which we can write down by inspection:

Eq. 2.65 $$\frac{d}{d\tau_\nu}\left(I_\nu e^{\tau_\nu}\right) = S_\nu e^{\tau_\nu}$$

And this can now be integrated from 0 to τ_ν :

Eq. 2.66 $$\left[I_\nu e^{\tau'_\nu}\right|_0^{\tau_\nu} = \int_0^{\tau_\nu} S_\nu(\tau'_\nu) e^{\tau'_\nu} d\tau'_\nu$$

Evaluating the left side gives the following:

Eq. 2.67 $$I_\nu(\tau_\nu)e^{\tau_\nu} - I_\nu(0) = \int_0^{\tau_\nu} S_\nu(\tau'_\nu) e^{\tau'_\nu} d\tau'_\nu$$

And finally, dividing by e^{τ_ν} gives us **the formal solution to the transfer equation for normal rays**:

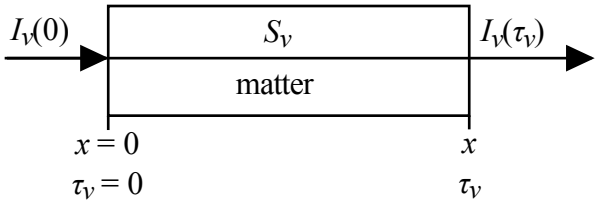
Eq. 2.68 $$I_\nu(\tau_\nu) = I_\nu(0)e^{-\tau_\nu} + \int_0^{\tau_\nu} S_\nu(\tau'_\nu) e^{-(\tau_\nu - \tau'_\nu)} d\tau'_\nu$$

The different pieces of this formal solution can be interpreted as follows:

- $I_\nu(0)\, e^{-\tau_\nu}$ is the original intensity, decreasing with τ_ν due to absorption.
- $\int_0^{\tau_\nu} S_\nu(\tau'_\nu) e^{-(\tau_\nu - \tau'_\nu)} d\tau'_\nu$ is the integrated source, increasing with τ'_ν (due to emission) and decreasing with τ_ν (due to absorption).

Figure 2.7 shows the coordinates associated with the ray as it passes through matter with source function S_ν . Without more knowledge of the source function, this solution cannot be further simplified.

Figure 2.7: Coordinates used in the formal solution to the transfer equation (for normal rays)



Let us now consider a simple case in which the source function is *constant* throughout the matter – that is, $S_\nu(\tau_\nu) = S_\nu$. In this case, S_ν can be brought outside of the integral and the integration can proceed as follows:

Eq. 2.69 $$I_\nu(\tau_\nu) = I_\nu(0)e^{-\tau_\nu} + S_\nu \int_0^{\tau_\nu} e^{-(\tau_\nu - \tau'_\nu)} d\tau'_\nu = I_\nu(0)e^{-\tau_\nu} + S_\nu \left(e^{-(\tau_\nu - \tau'_\nu)} \Big|_0^{\tau_\nu} \right.$$

Eq. 2.70 $$I_\nu(\tau_\nu) = I_\nu(0)\, e^{-\tau_\nu} + S_\nu(1 - e^{-\tau_\nu})$$

This result says that after passing through matter with an optical depth of τ_ν , the intensity of a ray will be a combination of the original intensity and the source function of the matter. Which of these two makes the greater contribution depends on the value of τ_ν . This can easily be seen by considering the behavior of the two exponential functions in Equation 2.70. Figure 2.8a shows the variation of the initial intensity term (where $x = \tau_\nu$) while Figure 2.8b shows the variation of the source function term.

Figure 2.8a: Graph of $y = e^{-x}$

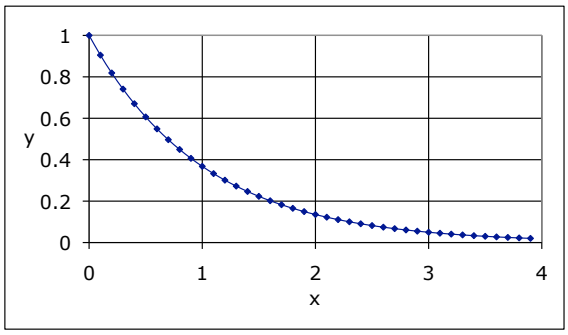
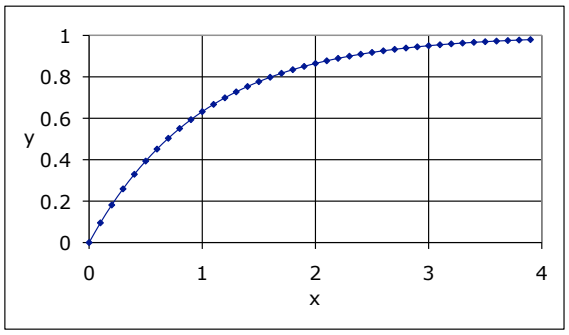


Figure 2.8b: Graph of $y = 1 - e^{-x}$



As τ_ν increases, the beam is affected less by the initial intensity and more by the source function of the medium. An example of this is found in the observations of a star through a cloud in the atmosphere. If the cloud is relatively thin – perhaps just haze – the star can still be seen with only minor attenuation. However, a thicker cloud not only blocks out the star better, making it harder to see, but also becomes more visible itself as its source function makes a greater contribution to the ray.

Note that at low optical depths ($\tau_\nu << 1$), both curves in Figure 2.8 become nearly linear. This is because we can use the series expansion for e^{-x} to approximate this function as a linear equation:

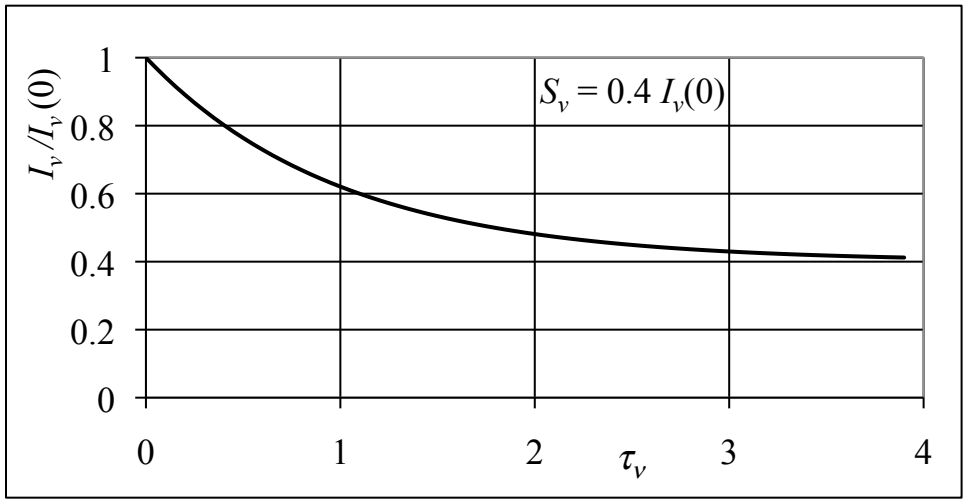
Eq. 2.71 $$e^{-x} = 1 - x + \frac{x^2}{2} - \ldots \approx 1 - x \quad \text{and} \quad 1 - e^{-x} \approx x \quad \text{for } x << 1$$

Then Equation 2.70 is further simplified:

Eq. 2.72 $$I_\nu(\tau_\nu) \approx I_\nu(0)(1 - \tau_\nu) + S_\nu \tau_\nu \qquad (\text{for } \tau_\nu << 1)$$

For low values of τ_ν , the source function will make a negligible contribution, and the intensity will be primarily due to the star. On the other hand, if the cloud is optically thick ($\tau_\nu >> 1$), then $e^{-\tau_\nu} \Rightarrow 0$, $1 - e^{-\tau_\nu} \Rightarrow 1$, and $I_\nu(\tau_\nu) \Rightarrow S_\nu$, meaning the radiation we see comes mostly from the cloud. In general, the star will be dimmed by the cloud, but the intensity will not drop below the source function of the cloud, as shown in Figure 2.9.

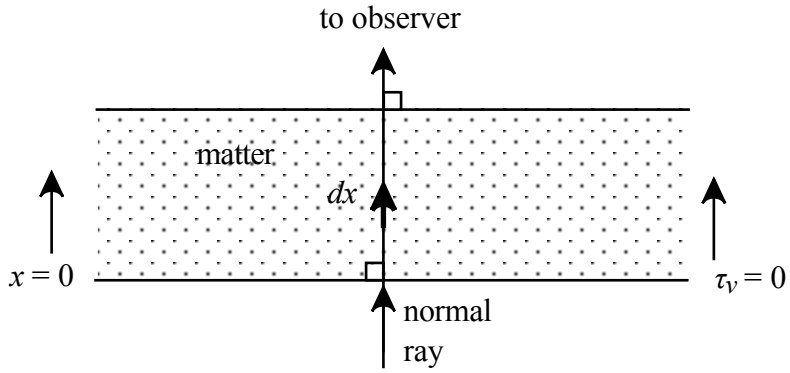
Figure 2.9: The intensity of a star as seen through a cloud with a source function that is 40% of the star's intensity



The Transfer Equation for Oblique Rays

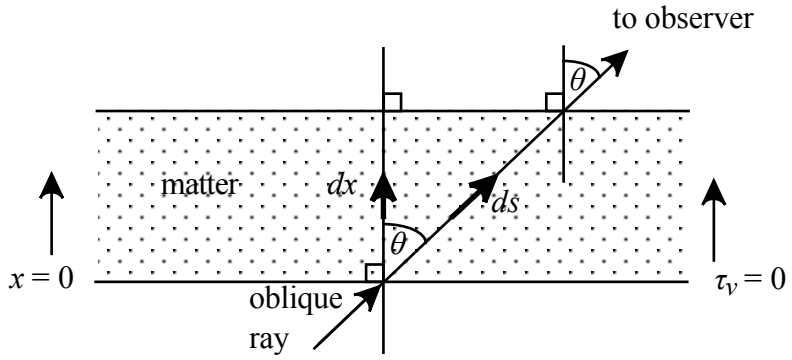
So far this equation of transfer applies only to radiation that is *normal* to the surface of the matter, as shown in Figure 2.10. The ray travels through the matter along the path *dx*.

Figure 2.10: Geometry for a normal ray



For the case of a ray that is *not* normal to the surface, an oblique path must be used, as shown in Figure 2.11.

Figure 2.11: Geometry for an oblique ray



To obtain the proper transfer equation, we must replace the **normal path length *dx*** in our existing equation by the **oblique path length *ds*** (= sec θ *dx*). Then

Eq. 2.35 $$\frac{dI_\nu}{dx} = -\kappa_\nu \rho I_\nu + j_\nu \rho$$ becomes

Eq. 2.73 $$\frac{dI_\nu}{ds} = -\kappa_\nu \rho I_\nu + j_\nu \rho \ .$$

Substituting *ds* = sec θ *dx* allows us to write the transfer equation for an oblique ray in terms of the normal coordinate *x*:

Eq. 2.74 $$\cos\theta \frac{dI_\nu}{dx} = -\kappa_\nu \rho I_\nu + j_\nu \rho$$

In solving the previous transfer equation (Equation 2.35), we made use of the optical depth, which depends on the path length. However, using the oblique path length *ds* will produce a different optical depth from that obtained using the normal path length *dx*, which is shorter. To solve this problem, we will work with *normal* coordinates and will define and use a *normal optical depth* even though the ray may not be normal.

Normal Optical Depth

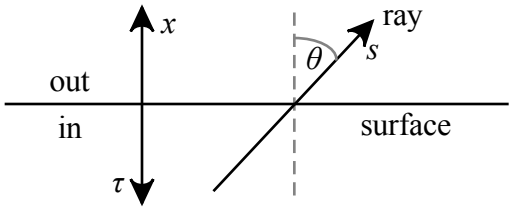
Our normal coordinate is *x*, and our **normal optical depth** may be defined by Equations 2.75:

Eq. 2.75a $$d\tau_\nu(x) \equiv \pm \kappa_\nu \rho \, dx$$

Eq. 2.75b $$\tau_\nu(x) = \pm \int_0^x \kappa_\nu \rho \, dx'$$

Here the '+' sign indicates that *x* and τ increase in the same direction (as in Figures 2.10 and 2.11), while the '–' sign indicates that they increase in opposite directions, which is often more convenient. In most astronomical applications, the ray moves outward from the surface toward the observer, but the observer looks backward along the ray, into the surface. For this reason we will modify our coordinates as shown in Figure 2.12.

Figure 2.12: Coordinates for the transfer function for oblique rays



Thus, we will be using the negative form of the normal optical depth:

Eq. 2.76 $$d\tau_\nu(x) \equiv -\kappa_\nu \rho\, dx$$

With this choice, the transfer equation shown in Equation 2.74 can be written in terms of the optical depth as shown:

Eq. 2.77 $$\cos\theta \frac{dI_\nu}{d\tau_\nu} = +I_\nu - S_\nu$$

This form of the transfer equation will be most appropriate when the properties of the medium are functions of the normal coordinate x. An example of this can be found in the atmosphere of a radially symmetric star for which the vertical extent of the atmosphere is much less than the radius of the star; such a case will be called a **plane-parallel atmosphere**, to be discussed below. First we should briefly investigate the problem of an extended atmosphere, for which this constraint does not apply.

Extended Atmospheres

Consider a star with an **extended atmosphere** – one for which the thickness of the atmosphere is *not* small compared to the radius of the star. The reason for this distinction is simple; for an extended atmosphere, a ray traversing the atmosphere enters and leaves at different angles θ (Figure 2.13a), while for a plane-parallel atmosphere, which is relatively thin compared to the star's radius, these angles are essentially the same (Figure 2.13b).

Figure 2.13a: Extended atmosphere

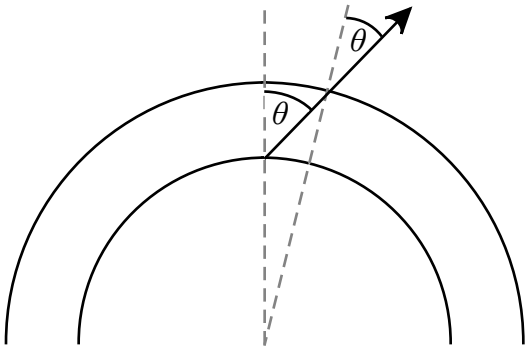
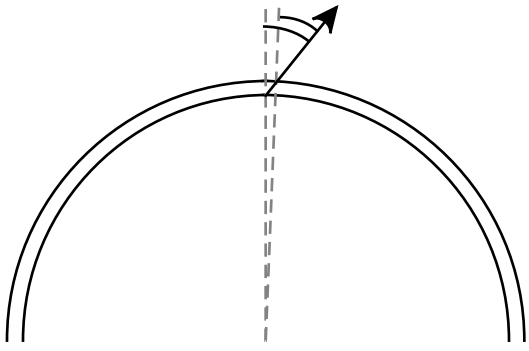


Figure 2.13b: Plane-parallel atmosphere



We will begin with the transfer equation of Equation 2.73, rewritten as follows:

Eq. 2.78 $$\frac{dI_\nu}{\kappa_\nu \rho\, ds} = -I_\nu + S_\nu$$ (The ray travels in the +s direction.)

For this problem, we will use spherical polar coordinates as shown in Figure 2.14a; these are oriented such that the *z*-axis (the polar axis) is the line of sight along the ray, pointing toward the observer; and the position vector $\vec{r}$ represents the normal to the surface (the stellar radius).

Figure 2.14a: Spherical polar coordinates

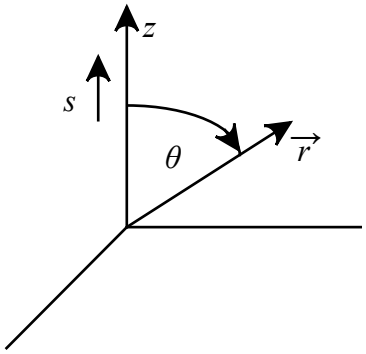
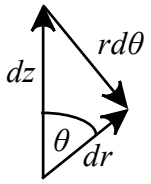


Figure 2.14b: Components of the ray



The differential ray vector *dz* can be resolved into components that are parallel (*rdθ*) and perpendicular (*dr*) to the surface, as shown in Figure 2.14b. The magnitudes of these components are $dr = dz \cos\theta$ and $rd\theta = -dz \sin\theta$, which lead to the derivatives given by Equations 2.79:

Eq. 2.79a $$\frac{dr}{dz} = \cos\theta$$ Eq. 2.79b $$\frac{d\theta}{dz} = -\frac{\sin\theta}{r}$$

The transfer equation (Equation 2.78) written in spherical polar coordinates is then as follows:

Eq. 2.80 $$\frac{1}{\kappa_\nu \rho}\frac{dI_\nu}{ds} = \frac{1}{\kappa_\nu \rho}\frac{dI_\nu}{dz} = \frac{1}{\kappa_\nu \rho}\left[\frac{\partial I_\nu}{\partial r}\frac{dr}{dz} + \frac{\partial I_\nu}{\partial \theta}\frac{d\theta}{dz}\right] = -I_\nu + S_\nu$$

Inserting the derivatives from Equations 2.79 yields the following result:

Eq. 2.81 $$\frac{1}{\kappa_\nu \rho}\left[\cos\theta\frac{\partial I_\nu}{\partial r} + \frac{\sin\theta}{r}\frac{\partial I_\nu}{\partial \theta}\right] = -I_\nu + S_\nu$$

This form of the transfer equation is used for extended atmospheres and for stellar interiors, where the angle between a particular ray and the normal is not constant with radius. We will ignore such cases and focus on relatively thin atmospheres, for which $\Delta R_{atmos}/R_* << 1$ and the ray enters and leaves the atmosphere at essentially the same angle. These are the conditions of a plane-parallel atmosphere. For such an atmosphere, because $d\theta/dz \approx 0$, the transfer equation (Equation 2.81) reduces to the following:

Eq. 2.82 $$\cos\theta\frac{dI_\nu}{\kappa_\nu \rho\, dr} = -I_\nu + S_\nu$$

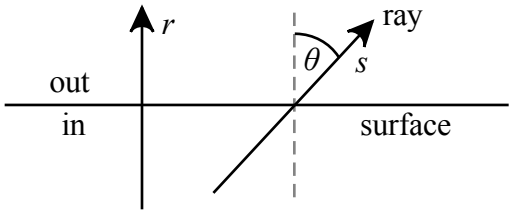
(An equivalent equation was previously derived (Equation 2.74), with *dx* in place of *dr*.)

Plane-Parallel Atmospheres

These cases are called plane-parallel atmospheres because the top and the bottom of the atmosphere are assumed to be parallel planes, even though the star is a sphere. (The approximation is good if the above condition $\Delta R_{atmos} << R_*$ is met.) We will now manipulate Equation 2.82 to obtain one that has the coordinates we need.

The current coordinates for Equation 2.82 are those shown in Figure 2.15. The outward normal coordinate is *r*, the stellar radius, and *s* is the coordinate outward along the ray.

Figure 2.15: Transfer equation (polar) coordinates



It will be convenient to also have an inward-directed coordinate, opposite *r* ; let us redefine our former outward *x*-coordinate as an *inward* normal coordinate such that $dx = -dr$. With this substitution, Equation 2.82 becomes as follows:

Eq. 2.83 $$\cos\theta \frac{dI_\nu}{\kappa_\nu \rho\, dx} = I_\nu - S_\nu$$ (Note the sign changes.)

Next we will introduce an optical depth. Because we want the ray moving outward but the optical depth increasing inwards, we could define $\tau_\nu(s)$ – the optical depth along the ray – by Equation 2.84:

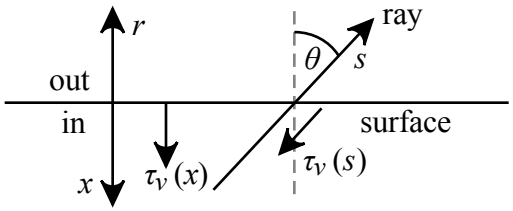
Eq. 2.84 $$d\tau_\nu(s) \equiv -\kappa_\nu \rho\, ds$$

But as noted above, it is more convenient to use a normal optical depth, which we can obtain by noting that $ds = -\sec\theta\, dx$, giving the following:

Eq. 2.85 $$d\tau_\nu(x) \equiv +\kappa_\nu \rho\, dx \qquad (\text{or } d\tau_\nu(r) \equiv -\kappa_\nu \rho\, dr\ )$$

Figure 2.16 shows a modified diagram of the coordinates used, including the optical depths.

Figure 2.16: Modified transfer equation (polar) coordinates



Writing the transfer equation in terms of the optical depth $\tau_\nu(s)$ results in the following form:

Eq. 2.86 $$\cos\theta \frac{dI_\nu}{d\tau_\nu} = I_\nu - S_\nu$$

This can now be solved to give the intensity of the ray as a function of optical depth.

Solution of the Transfer Equation

To aid in the solution of this differential equation, let us make a simple substitution, setting $\cos\theta$ equal to μ. This changes the transfer equation to the following:

Eq. 2.87 $$\mu\frac{dI_\nu}{d\tau_\nu} - I_\nu = -S_\nu$$

This can be solved by multiplying both sides of the equation by $e^{-\tau_\nu/\mu}$ as shown:

Eq. 2.88 $$e^{-\tau_\nu/\mu}\frac{dI_\nu}{d\left(\tau_\nu/\mu\right)} - e^{-\tau_\nu/\mu}I_\nu = -e^{-\tau_\nu/\mu}S_\nu$$

It will be immediately recognized that the left side of the equation is an exact differential:

Eq. 2.89 $$\frac{d}{d\left(\tau_\nu/\mu\right)}\left[e^{-\tau_\nu/\mu}I_\nu\right] = -e^{-\tau_\nu/\mu}S_\nu$$

This can be integrated to give the following equation (where t_ν is a dummy integration variable):

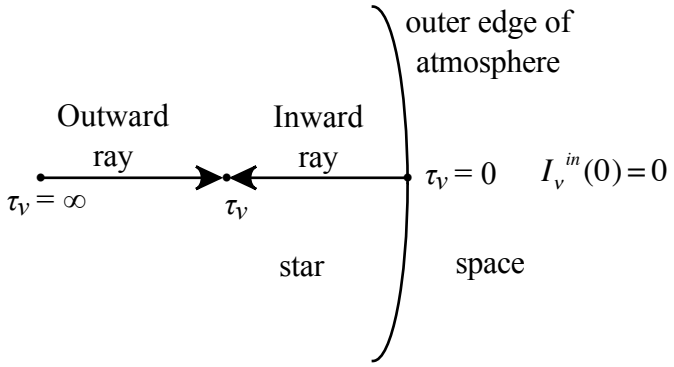
Eq. 2.90 $$e^{-\frac{\tau_\nu}{\mu}}I_\nu = -\int_C^{\frac{\tau_\nu}{\mu}} S_\nu e^{-\frac{t_\nu}{\mu}} d\left(t_\nu/\mu\right)$$

And this can be rearranged to yield the solution:

Eq. 2.91 $$I_\nu(\tau_\nu) = -\int_C^{\tau_\nu} S_\nu(t_\nu)e^{-(t_\nu-\tau_\nu)/\mu}dt_\nu/\mu$$

The integration limit C depends on the boundary conditions of the problem, and this will depend on whether the radiation is directed inward or outward.

Figure 2.17: Coordinates associated with the solution to the transfer equation for an isolated star



Consider an isolated star, surrounded by space, with essentially no radiation incident on it, as shown if Figure 2.17. For inward-directed rays inside the atmosphere, the integration will then proceed from the outer edge of the atmosphere – in practice, a not-very-well-defined location –

to the position in question within the atmosphere at optical depth τ_ν and intensity $I_\nu(\tau_\nu)$. The appropriate value of the τ_ν integration limit C is 0. With this substitution, the intensity of inward rays is as follows:

Eq. 2.92 $$I_\nu^{in}(\tau_\nu) = -\int_0^{\tau_\nu} S_\nu(t_\nu) e^{-(t_\nu - \tau_\nu)/\mu} dt_\nu/\mu$$

For outward-directed rays in a star, photons may originated at greater optical depths, as deep as $\tau_\nu = \infty$, as shown in Figure 2.17; therefore $C = \infty$ for this integration.

Eq. 2.93 $$I_\nu^{out}(\tau_\nu) = -\int_\infty^{\tau_\nu} S_\nu(t_\nu) e^{-(t_\nu - \tau_\nu)/\mu} dt_\nu/\mu = \int_{\tau_\nu}^\infty S_\nu(t_\nu) e^{-(t_\nu - \tau_\nu)/\mu} dt_\nu/\mu$$

The total intensity at optical depth τ_ν is then the sum of the inward and outward intensities:

Eq. 2.94 $$I_\nu(\tau_\nu) = I_\nu^{out}(\tau_\nu) + I_\nu^{in}(\tau_\nu) = \int_{\tau_\nu}^\infty S_\nu(t_\nu) e^{-(t_\nu - \tau_\nu)/\mu} dt_\nu/\mu - \int_0^{\tau_\nu} S_\nu(t_\nu) e^{-(t_\nu - \tau_\nu)/\mu} dt_\nu/\mu$$

Although measuring the intensity at different points within a star's atmosphere would be very difficult, we should be able to measure the value at the top of the atmosphere, where $\tau_\nu = 0$, and therefore $I_\nu^{in}(0) = 0$. This leads directly to the following expression for the total intensity, which is just the outward intensity:

Eq. 2.95 $$I_\nu(0) = I_\nu^{out}(0) = \int_0^\infty S_\nu(t_\nu) e^{-t_\nu/\mu} dt_\nu/\mu$$

Of course, this is only useful for stars for which different values of μ (= cos θ) can be distinguished – such as the Sun. Most stars appear to us only as points, for which only the flux can be measured.

Given our expression for the specific intensity (Equation 2.94), if we know the source function's dependence on the optical depth, we can determine not only the intensity, but also the mean density, flux, radiation pressure, etc. Let us see what form such equations will take.

From before (Equation 2.7), we have an expression for the flux, which can now be rewritten using the standard substitution (μ = cos θ and $d\mu$ = –sin θ):

Eq. 2.96 $$F_\nu = \int_{4\pi} I_\nu \cos\theta \, d\omega = 2\pi \int_0^\pi I_\nu \cos\theta \sin\theta \, d\theta = 2\pi \int_{-1}^1 I_\nu \, \mu \, d\mu$$

We can separate the integral into outward and inward components:

Eq. 2.97 $$F_\nu = 2\pi \int_0^1 I_\nu^{out} \, \mu \, d\mu + 2\pi \int_{-1}^0 I_\nu^{in} \, \mu \, d\mu$$

Now we can insert our expressions for the intensity from Equation 2.94 into the appropriate integrals:

Eq. 2.98 $$F_\nu(\tau_\nu) = 2\pi \int_0^1 \int_{\tau_\nu}^\infty S_\nu(t_\nu) e^{-(t_\nu - \tau_\nu)/\mu} dt_\nu \, d\mu - 2\pi \int_{-1}^0 \int_0^{\tau_\nu} S_\nu(t_\nu) e^{-(t_\nu - \tau_\nu)/\mu} dt_\nu \, d\mu$$

If we make our standard assumption that the source function is isotropic, then S_ν can be removed from the μ integral (but not the t integral) and the order of integration reversed.

Eq. 2.99 $$F_\nu(\tau_\nu) = 2\pi \int_{\tau_\nu}^{\infty} S_\nu(t_\nu) \int_0^1 e^{-(t_\nu - \tau_\nu)/\mu} d\mu \, dt_\nu - 2\pi \int_0^{\tau_\nu} S_\nu(t_\nu) \int_{-1}^0 e^{-(t_\nu - \tau_\nu)/\mu} d\mu \, dt_\nu$$

In order to make this expression easier to write, it is customary to make the following substitutions: let $\mu \equiv 1/w \;\Rightarrow\; d\mu = -\, dw/w^2$, and let $x \equiv t_\nu - \tau_\nu$. (Note: This is *not* our inward normal coordinate x.) Then we can rewrite the μ integral as follows:

Eq. 2.100 $$\int_0^1 e^{-(t_\nu - \tau_\nu)/\mu} d\mu = \int_\infty^1 -\frac{e^{-wx}}{w^2} dw = \int_1^\infty \frac{e^{-wx}}{w^2} dw$$

This is the well-known **exponential integral** $E_2(x)$, of the following general form:

Eq. 2.101 $$E_n(x) \equiv \int_1^\infty \frac{e^{-xw}}{w^n} dw$$

Exponential Integrals

The **exponential integrals $E_n(x)$** are monotonically diminishing functions of x that all behave the same way for large values of x (Gray 1976). They serve as extinction factors in the solutions to the equations of radiative transfer.

General form: $$E_n(x) = \int_1^\infty \frac{e^{-xw}}{w^n} dw$$

Special cases:

$$E_n(0) = \int_1^\infty w^{-n} \, dw$$

If $n \neq 0, 1$ $$E_n(0) = \frac{1}{1-n} w^{1-n} \Big|_1^\infty = \frac{\infty^{1-n}}{1-n} - \frac{1}{1-n} = \frac{1}{n-1}$$

If $n = 0$ $$E_0(x) = \int_1^\infty e^{-xw} \, dw = \frac{e^{-xw}}{-x} \Big|_1^\infty = 0 - \frac{e^{-x}}{-x} = \frac{e^{-x}}{x}$$

and $$E_0(0) = \frac{1}{0} \Rightarrow \infty$$

If $n = 1$ $$E_1(0) = \int_1^\infty \frac{dw}{w} = \ln w \Big|_1^\infty = \ln\infty - 0 = \infty$$

Recursion relations: $$\frac{dE_n(x)}{dx} = \int_1^\infty \frac{1}{w^n} \frac{d}{dx} e^{-xw} dw = -\int_1^\infty \frac{e^{-xw}}{w^{n-1}} dw = -E_{n-1}(x)$$

and $$nE_{n+1}(x) = e^{-x} - xE_n(x)$$

We can accomplish a similar simplification for the second μ integral in Equation 2.99 by letting $\mu \equiv -1/w \;\Rightarrow\; d\mu = dw/w^2$, and $x \equiv \tau_\nu - t_\nu$:

Eq. 2.102 $$\int_{-1}^0 e^{-(t_\nu - \tau_\nu)/\mu} d\mu = \int_1^\infty \frac{e^{-xw}}{w^2} dw$$

Then the flux can be written as follows:

Eq. 2.103 $$F_\nu(\tau_\nu) = 2\pi \int_{\tau_\nu}^{\infty} S_\nu(t_\nu) E_2(t_\nu - \tau_\nu) dt_\nu - 2\pi \int_0^{\tau_\nu} S_\nu(t_\nu) E_2(\tau_\nu - t_\nu) dt_\nu$$

The theoretical flux at the surface of the star ($\tau_\nu = 0$) is found from Equation 2.103:

Eq. 2.104 $$F_\nu(0) = 2\pi \int_0^{\infty} S_\nu(t_\nu) E_2(t_\nu) dt_\nu$$

Equations for the mean intensity and the radiation pressure integral can also be obtained in a similar fashion:

Eq. 2.105 $$J_\nu(\tau_\nu) = \frac{1}{2} \int_{\tau_\nu}^{\infty} S_\nu(t_\nu) E_1(t_\nu - \tau_\nu) dt_\nu + \frac{1}{2} \int_0^{\tau_\nu} S_\nu(t_\nu) E_1(\tau_\nu - t_\nu) dt_\nu$$

Eq. 2.106 $$K_\nu(\tau_\nu) = \frac{1}{2} \int_{\tau_\nu}^{\infty} S_\nu(t_\nu) E_3(t_\nu - \tau_\nu) dt_\nu + \frac{1}{2} \int_0^{\tau_\nu} S_\nu(t_\nu) E_3(\tau_\nu - t_\nu) dt_\nu$$

We are now ready to apply our basic radiative transfer results to the problem of radiation passing through the stellar atmosphere, a region of the star also known as the **photosphere**.

Radiative Equilibrium in the Photosphere

A star generates energy in its optically thick interior and transports this energy to the stellar surface, where photons carry the energy into space. In the stellar photosphere, the gas changes from optically thick to optically thin, meaning that *most* of the photons will pass through this region, but some will not. Those photons that *are* absorbed will deposit their energy in the photospheric gas, but this energy will be re-emitted in the form of new photons. The net effect of the photosphere is then to *transmit* energy; it serves as neither a source nor a sink of photons. Such a gas is said to be in **radiative equilibrium**.

With a photosphere in radiative equilibrium, energy flows into the base of the atmosphere at the same rate that it flows out of the top; in other words, the luminosity is constant throughout the photosphere. Of course, luminosity is the product of flux and surface area, and because we are using a plane-parallel atmosphere, the top and bottom of the atmosphere should have the same surface area. Therefore, with constant luminosity and constant area throughout the atmosphere, the flux should also be constant – a value we will denote as F_0. (It is assumed that the sole mechanism of energy transport in the atmosphere is radiation – that convection and conduction play no significant role.)

As noted above, we are allowing absorption to occur, followed by emission, processes that may involve photons of different frequencies. Therefore the constant flux condition is applicable to the integrated flux, rather than F_ν:

Eq. 2.107 $$F_0 = \int_0^{\infty} F_\nu d\nu = \text{constant}$$

Also, we can integrate the transfer equation (derived from Equations 2.85 and 2.86) over all angles, as follows:

Eq. 2.108 $$\cos\theta \frac{dI_\nu}{dx} = \kappa_\nu \rho I_\nu - \kappa_\nu \rho S_\nu$$

Eq. 2.109 $$\frac{d}{dx}\int_{4\pi} I_\nu \cos\theta\, d\omega = \int_{4\pi} \kappa_\nu \rho\, I_\nu\, d\omega - \int_{4\pi} \kappa_\nu \rho\, S_\nu\, d\omega$$

Now let us assume that κ_ν , ρ , and S_ν are independent of ω; then they can be removed from the integrals, turning each integral into a recognizable expression:

Eq. 2.110 $$\frac{d}{dx}\int_{4\pi} I_\nu \cos\theta\, d\omega = \kappa_\nu \rho \int_{4\pi} I_\nu\, d\omega - \kappa_\nu \rho\, S_\nu \int_{4\pi} d\omega$$

Eq. 2.111 $$\frac{dF_\nu}{dx} = 4\pi\kappa_\nu \rho\, J_\nu - 4\pi\kappa_\nu \rho\, S_\nu$$

We then integrate over frequency:

Eq. 2.112 $$\frac{d}{dx}\int_0^\infty F_\nu d\nu = 4\pi\rho \int_0^\infty \kappa_\nu J_\nu d\nu - 4\pi\rho \int_0^\infty \kappa_\nu S_\nu d\nu$$

The left side of the equation is zero because the integrated flux is constant with x. This leaves the following result:

Eq. 2.113 $$\int_0^\infty \kappa_\nu J_\nu d\nu = \int_0^\infty \kappa_\nu S_\nu d\nu$$

A third result can be obtained by returning to the transfer equation (Equation 2.108) and multiplying both sides by cos θ:

Eq. 2.114 $$\cos^2\theta \frac{dI_\nu}{dx} = \kappa_\nu \rho\, I_\nu \cos\theta - \kappa_\nu \rho\, S_\nu \cos\theta$$

This equation is then integrated over all angles, with the same assumptions about κ_ν , ρ , and S_ν being independent of ω:

Eq. 2.115 $$\frac{d}{dx}\int_{4\pi} I_\nu \cos^2\theta\, d\omega = \kappa_\nu \rho \int_{4\pi} I_\nu \cos\theta\, d\omega - \kappa_\nu \rho\, S_\nu \int_{4\pi} \cos\theta\, d\omega$$

Eq. 2.116 $$4\pi \frac{dK_\nu}{dx} = \kappa_\nu \rho\, F_\nu - 0 \quad \Rightarrow \quad \frac{dK_\nu}{d\tau_\nu} = \frac{F_\nu}{4\pi}$$

And integrating this expression over frequency yields the following:

Eq. 2.117 $$\int_0^\infty \frac{dK_\nu}{d\tau_\nu} d\nu = \frac{F_0}{4\pi}$$

The condition of radiative equilibrium gives the following three results:

Eq. 2.107 $$F_0 = \int_0^\infty F_\nu d\nu = \text{constant}$$

Eq. 2.113 $$\int_0^\infty \kappa_\nu J_\nu d\nu = \int_0^\infty \kappa_\nu S_\nu d\nu$$

Eq. 2.117 $$\int_0^\infty \frac{dK_\nu}{d\tau_\nu} d\nu = \frac{F_0}{4\pi}$$

The flux constant F_0 is usually expressed in terms of a temperature:

Eq. 2.118 $F_0 \equiv \sigma T_e^4$

Here σ is the Stefan-Boltzmann constant (≈ 5.670e–5), and T_e is the effective temperature, which is the temperature of a blackbody that has the same integrated flux as the star. The link is as follows:

A blackbody has an integrated intensity $I = B(T) = \sigma T^4/\pi$ (Equation 2.52), and the outward flux at the surface of a star (from an isotropic field) is $F = \pi I$ (Equation 2.10). Combining these two expressions gives Equation 2.118.

Let us now make use of a rather unrealistic example.

The Gray Case

Real stars are very complicated; one of the main complications is that photons of different frequencies interact differently with stellar matter – which is why we have been carrying subscripts on our absorption coefficients, intensities, and such. Radiative transfer problems would be easier to solve if this were not the case.

As an illustration, let us assume that absorption has *no* frequency dependence – that $\kappa_\nu = \kappa$. There will then be no need to keep track of different frequencies, and we can write $I = \int_0^\infty I_\nu d\nu$, etc. This is called the **gray case** (because with no frequency differentiation, there would be no color difference – only shades of gray depending on the magnitude of the intensity). The transfer equation (Equation 2.108) then is written as follows:

Eq. 2.119 $$\cos\theta \frac{d}{dx}\int_0^\infty I_\nu d\nu = \kappa\rho\int_0^\infty I_\nu d\nu - \kappa\rho\int_0^\infty S_\nu d\nu$$ or

Eq. 2.120 $$\cos\theta \frac{dI}{dx} = \kappa\rho I - \kappa\rho S$$ or

Eq. 2.121 $$\cos\theta \frac{dI}{d\tau} = I - S$$

Similarly, the radiative equilibrium conditions become much simpler:

Eq. 2.122 $F = F_0 = \text{constant}$

Eq. 2.123 $J = S$

Eq. 2.124 $$\frac{dK}{d\tau} = \frac{F_0}{4\pi}$$ (a constant = H_0)

The first of these says that the flux is constant. The second says that the source function is equal to the mean intensity; determination of $J(\tau)$ will give us $S(\tau)$. The third equation can be easily integrated:

Eq. 2.125 $$K(\tau) = \frac{F_0}{4\pi}\tau + \text{constant}$$ or $K(\tau) = H_0\tau + \text{constant}$

With one more simplifying assumption, we will be able to determine both the intensity and the source function for the gray atmosphere.

The Eddington Approximation

We will not be able to proceed very far in our quest without having to integrate over solid angle. Previously we have done so by assuming the radiation field was isotropic, but this will not always be an appropriate assumption. We need a solid angle distribution that is simple enough to be integrable and yet still somewhat realistic for a stellar atmosphere. This is provided by the **Eddington approximation**.

In an isotropic radiation field, the intensity is constant over the whole sphere (4π steradians); that is, at a given point in the atmosphere, at some particular optical depth, the intensity is the same in all directions (but it still varies with τ). The Eddington approximation uses a radiation field that is constant over the outward hemisphere and constant over the inward hemisphere, but the two constants are different. At any given optical depth τ, the intensity takes on one of two values, depending on the direction of the ray.

For outward rays, for which $\theta = 0 \rightarrow \pi/2$, $I(\tau) = I^{+}(\tau)$.

For inward rays, for which $\theta = \pi/2 \rightarrow \pi$, $I(\tau) = I^{-}(\tau)$.

From this starting point, we can determine the mean intensity (using Equation 2.2):

Eq. 2.126 $$J(\tau) = \frac{1}{4\pi}\int_{4\pi} I\, d\omega = \frac{1}{2}\int_0^1 I^{+}\, d\mu + \frac{1}{2}\int_{-1}^0 I^{-}\, d\mu = \frac{1}{2} I^{+}\int_0^1 d\mu + \frac{1}{2} I^{-}\int_{-1}^0 d\mu$$

Eq. 2.127 $$J(\tau) = {}^1\!/_2\,[I^{+} + I^{-}]$$

The flux can be found in a similar fashion:

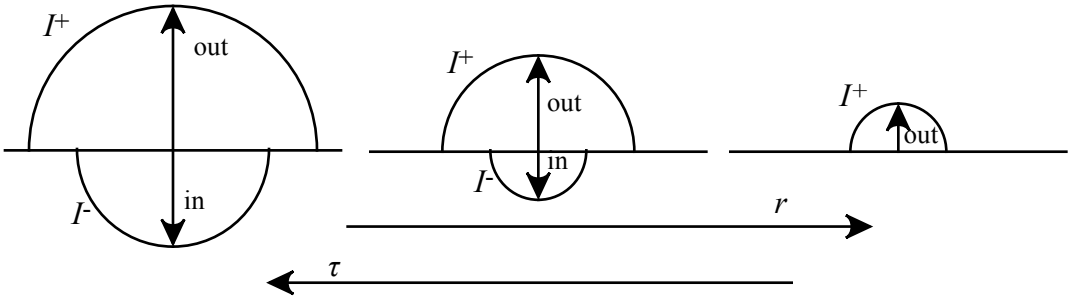
Eq. 2.128 $$F(\tau) = \int_{4\pi} I\,\mu\, d\omega = 2\pi\int_0^1 I^{+}\,\mu\, d\mu + 2\pi\int_{-1}^0 I^{-}\,\mu\, d\mu = 2\pi\left[I^{+}\int_0^1 \mu\, d\mu + I^{-}\int_{-1}^0 \mu\, d\mu\right]$$

Eq. 2.129 $$F(\tau) = 2\pi\left[I^{+}\left(\frac{1}{2}\right) + I^{-}\left(\frac{-1}{2}\right)\right] = \pi\,[I^{+}(\tau) - I^{-}(\tau)] = \text{constant} = F_0$$

This result indicates that the difference between the outward intensity and the inward intensity is a constant. Even though each of these intensities varies with optical depth, the difference between them is the same throughout the atmosphere.

We can represent this progression in a series of diagrams, in which semicircles indicate the inward and outward radiation, with the radius of each semicircle proportional to the intensity. In moving from left to right in Figure 2.18, the stellar radius increases, indicating a higher level in the atmosphere and consequently a lower optical depth. As the optical depth decreases, so do both I^{+} and I^{-}, but the difference between them is constant.

Figure 2.18: Illustration of the Eddington approximation



The last sketch on the right represents the top of the atmosphere, where $\tau = 0$ and I^- drops to 0. At this point, the flux is equal to π times the outward intensity:

Eq. 2.130 $$\pi I^+(0) = F_0$$

The Eddington approximation also allows the radiation pressure integral to be solved:

Eq. 2.131 $$K(\tau) = \frac{1}{4\pi}\int_{4\pi} I\,\mu^2 d\omega = \frac{1}{2}\int_0^1 I^+\mu^2 d\mu + \frac{1}{2}\int_{-1}^0 I^-\mu^2 d\mu = \frac{1}{2}I^+\int_0^1 \mu^2 d\mu + \frac{1}{2}I^-\int_{-1}^0 \mu^2 d\mu$$

Eq. 2.132 $$K(\tau) = \frac{1}{2}\left[I^+\left(\frac{\mu^3}{3}\Big|_0^1\right) + I^-\left(\frac{\mu^3}{3}\Big|_{-1}^0\right)\right] = \frac{1}{6}\left[I^+ + I^-\right] = \frac{1}{3}J(\tau)$$

Solution of the differential equation 2.124 resulted in a constant, which can be evaluated by examining boundary conditions for $K(\tau)$. We can obtain the necessary boundary condition for the top of the atmosphere as follows:

As already noted, we have $I^-(0) = 0$, which leads to $J(0) = {}^1\!/_2\,[I^+(0) + I^-(0)] = {}^1\!/_2\,I^+(0)$. Applying Equation 2.130, we can link the mean intensity to the flux:

Eq. 2.133 $$J(0) = \frac{1}{2}I^+(0) = \frac{1}{2}\frac{F_0}{\pi}$$

Then evaluating Equation 2.132 at the top of the atmosphere gives a value for $K(0)$:

Eq. 2.134 $$K(0) = \frac{1}{3}J(0) = \frac{F_0}{6\pi}$$

We can also evaluate $K(0)$ using Equation 2.125:

Eq. 2.125 $$K(\tau) = \frac{F_0}{4\pi}\tau + \text{constant} \quad \Rightarrow \quad K(0) = \text{constant} = \frac{F_0}{6\pi}$$

The functions $K(\tau)$ and $J(\tau)$ can then be written explicitly:

Eq. 2.135 $$K(\tau) = \frac{F_0}{4\pi}\left(\tau + \frac{2}{3}\right)$$

Eq. 2.136 $$J(\tau) = \frac{3F_0}{4\pi}\left(\tau + \frac{2}{3}\right) = S(\tau)$$

(The function $J(\tau)$ is equal to $S(\tau)$, as one of the conditions of radiative equilibrium.)

With this result, we have the source function $S(\tau) = {}^3/_{4\pi}\, F_0(\tau + {}^2/_3)$ for the gray case ($\kappa_\nu = \kappa$) using the Eddington approximation (I^+ and I^- each constant over a hemisphere). We can carry it one step further if we add the condition of LTE (local thermodynamic equilibrium), meaning the radiation is in equilibrium with the gas. In this case, the source function is given by the Planck function:

Eq. 2.137 $$S = B(T) = \frac{\sigma T^4}{\pi}$$

Inserting this result into Equation 2.136 gives a relation for the temperature as a function of optical depth:

Eq. 2.138 $$\frac{3F_0}{4\pi}\left(\tau + \frac{2}{3}\right) = \frac{\sigma T(\tau)^4}{\pi}$$

The flux can be written in terms of the effective temperature ($F_0 \equiv \sigma T_e^4$), and this allows us to formulate equation 2.138 using only temperatures:

Eq. 2.139 $$T(\tau)^4 = \frac{3}{4}\left(\tau + \frac{2}{3}\right) T_e^4 = \left(\frac{1}{2} + \frac{3}{4}\tau\right) T_e^4$$

At what optical depth is the temperature equal to the effective temperature? Equation 2.139 gives $T(\tau) = T_e$ at $\tau = {}^2/_3$. Although this result has been derived for a particular model (the gray case, with the Eddington approximation), it is often used as an approximation for stellar atmospheres in general.

Using Equation 2.139, we can determine a temperature at the top of the atmosphere, at $\tau = 0$:

Eq. 2.140 $$T_0^4 \equiv T(0)^4 = {}^1/_2\, T_e^4$$

For the Sun, with an effective temperature of 5770 K, the value of T_0 is 4850 K. However, it should be noted that this temperature is not particularly meaningful or useful.

We can write the source function for the gray case with the Eddington approximation in terms of either the effective temperature or the flux:

Eq. 2.141 $$S(\tau) = \frac{\sigma T_e^4}{\pi}\left(\frac{1}{2} + \frac{3}{4}\tau\right) = \frac{F_0}{2\pi} + \frac{3F_0}{4\pi}\tau$$

In either case, it is clear that the source function is a linear function of the optical depth:

Eq. 2.142 $$S(\tau) = a + b\tau$$

Limb Darkening

We can insert this equation for the source function into our solution of the transfer equation to obtain predictions that can then be compared with observations. Equation 2.91 gives the intensity as a function of optical depth, but it is also dependent upon the angle of the ray, making $I(\tau)$ really $I(\tau, \mu)$. For observational purposes, we would be looking at the rays that emerge from

the top of the atmosphere, where $\tau = 0$, making $I(0, \mu)$ the function of interest; this would give us the variation in intensity across the disk of the star. Unfortunately, most stars occur only as points, for which intensity cannot be measured; but for the Sun, we can measure intensities across the disk and compare them with our predictions.

We may begin with Equation 2.95, which gives the intensity at the top of the atmosphere:

Eq. 2.95 $$I_\nu(0) = I_\nu^{out}(0) = \int_0^\infty S_\nu(t_\nu) e^{-t_\nu/\mu} dt_\nu/\mu$$

We now insert the source function (for the gray case, with the Eddington approximation):

Eq. 2.143 $$I(0,\mu) = \int_0^\infty (a+bt) e^{-t/\mu} dt/\mu = \int_0^\infty a\, e^{-t/\mu} dt/\mu + \mu \int_0^\infty b \frac{t}{\mu} e^{-t/\mu} dt/\mu$$

Substituting $x = \tau/\mu$ gives a simpler equation:

Eq. 2.144 $$I(0,\mu) = \int_0^\infty a\, e^{-x} dx + b\mu \int_0^\infty x\, e^{-x} dx$$

Performing the integrals yields an interesting result:

Eq. 2.145 $$I(0,\mu) = a\left(-e^{-x}\Big|_0^\infty + b\mu\left(-e^{-x}(x+1)\right)\Big|_0^\infty = a + b\mu\right.$$

Inserting the values for the constants a and b from Equation 2.141 gives the following:

Eq. 2.146 $$I(0,\mu) = a + b\mu = \frac{F_0}{2\pi} + \frac{3F_0}{4\pi}\mu$$

We may use this equation to predict the intensity at the surface of the Sun at different points across the disk. At the center of the disk, rays that reach us are emerging parallel to the normal with $\theta = 0°$, making μ (= cos θ) = 1:

Eq. 2.147 $$I(0,1) = \frac{F_0}{2\pi} + \frac{3F_0}{4\pi} = \frac{5F_0}{4\pi}$$

At the limb of the Sun, rays that reach us are emerging perpendicular to the normal with $\theta = 90°$, making $\mu = 0$:

Eq. 2.148 $$I(0,0) = \frac{F_0}{2\pi}$$

The intensity at the limb is clearly less than the intensity at the center of the disk, making the limb appear darker – a phenomenon known as **limb darkening**. It will be most convenient to remove the flux from the expression by dividing equation 2.146 by the maximum value, as limb darkening is a *relative* effect. This will produce the **normalized limb-darkening function**:

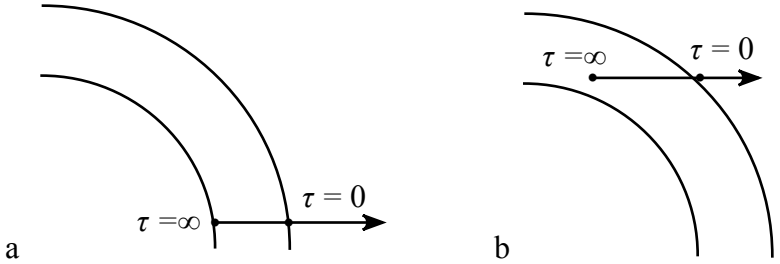
Eq. 2.149 $$\frac{I(0,\mu)}{I(0,1)} = \frac{2}{5} + \frac{3}{5}\mu$$

This function has a value of 1 at the center of the disk and 2/5 at the limb; obviously the mathematics confirms limb darkening. Is there an intuitive reason why it should occur?

Consider our line of sight through the star's atmosphere. We can see along this line only until an optical depth of infinity is reached; photons emitted from deeper in the star than this have no chance of traveling to us along this line. Figure 2.19a shows such a ray at the center of the disk, emitted along the normal towards us. The deepest point from which photons emerge is the base of the photosphere, marked $\tau = \infty$.

In Figure 2.19b, the ray emerges at an angle closer to 90° from the normal as it travels from the limb towards us. This ray's path through the atmosphere (from $\tau = \infty$ to $\tau = 0$) is longer because, being higher in the atmosphere, it passes through lower densities, and $\tau \approx \kappa\rho r$. For lower values of ρ, larger values of r will be needed to achieve the same τ.

Figure 2.19: Rays emerging from the center of the disk (a) and from the limb (b).



Also note that the origin ($\tau = \infty$) of the ray in Figure 2.19b is *higher* in the atmosphere, where the temperature – and thus, the source function – will be *lower*. This means that when looking at the limb, we are viewing photons that were emitted by gas with a lower source function compared to those coming from the center of the disk. Because the source function is less, the intensity will be lower and the limb will appear darker.

The Next Step

The problem we are preparing to solve using radiative transfer is the determination of the structure of the photosphere. This means that we want to find out how temperature (T), pressure (P), and density (ρ) of the gas vary with photospheric depth (x) and optical depth (τ). We now have some radiative transfer equations that provide these links; however, these equations involve quantities that depend on each other in an often complicated manner. For example, the optical depth is a function of density, opacity, and radius, while the opacity is a function of density, temperature, and composition. And we have just found that temperature is a function of optical depth; in short, everything is interrelated.

The key to this enterprise is the opacity (κ), which depends on the properties of matter. Therefore, we will need to examine atoms to see how they relate to radiation – to find out how absorption occurs. This will be done in the next chapter.

CHAPTER 3: Atomic Structure

In Chapter 2, we allowed for radiation to interact with matter, and we permitted the interaction to be dependent on the frequency of the radiation (although we found it simpler to ignore this dependency). In this chapter, we will investigate the nature of matter and how its structure makes it possible for photons and atoms to relate to each other.

One-Electron Atoms

At the start of the 20th century the structure of the atom was still unclear; the electron had just been discovered in 1897, but the proton would not be identified until 1918. With this minimal knowledge of the participant particles, physicists were hard pressed to explain the interactions between radiation and matter that had been observed for the previous few decades – interactions that were clearly frequency (or wavelength) dependent and varied from element to element. An empirical formula had been developed that predicted the wavelengths of hydrogen emission lines, but a suitable theoretical model of even this simplest atom was not available.

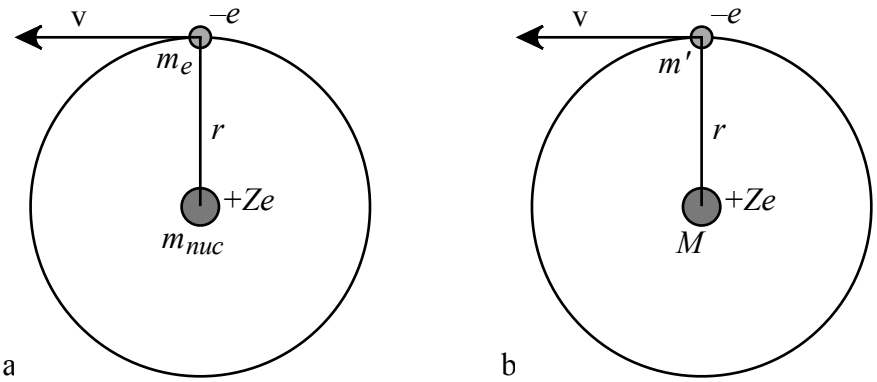
The Bohr Model

In 1913 Niels Bohr proposed a model of the hydrogen atom that filled this gap reasonably well. The **Bohr model** applies only to single-electron atoms and ions and does not predict all the observed spectral features, but it still is quite useful as a first approximation.

The Bohr model of the atom pictures the negatively charged electron (with mass m_e and charge q) moving at speed v in a circular orbit of radius r about a positively charged nucleus (with mass m_{nuc} and charge Q), as shown in Figure 3.1a. The two particles are held together by the electrostatic force between them, which is given by Coulomb's law: $F = Qq/r^2$.

We note immediately that this is another example of the two-body problem with a central force. In such problems, both particles will orbit about their common center of mass – an unnecessary complication we choose to avoid. We can do this by transforming to the equivalent reduced mass problem, in which the *stationary* nucleus, now with mass $M = m_{nuc} + m_e$, is orbited by the electron, now with the **reduced mass** $m' = m_e m_{nuc}/(m_e + m_{nuc})$, as shown in Figure 3.1b.

Figure 3.1: The Bohr model: (a) as a two-body problem; (b) as a reduced mass problem



The charges are unchanged by this transformation; we will assign the electron a charge of $q = -e$ (where e is the fundamental electron charge*), and the nuclear charge will be $Q = +Ze$ (where Z is the atomic number of the atom). The analysis then proceeds as follows.

The magnitude of the force between the nucleus and the electron is given by Coulomb's law:

Eq. 3.1 $$F = \left|\frac{Qq}{r^2}\right| = \left|\frac{(Ze)(-e)}{r^2}\right| = \frac{Ze^2}{r^2}$$

And the magnitude of the acceleration of a particle moving at speed v in a circular path of radius r is as follows:

Eq. 3.2 $$a = \frac{\mathrm{v}^2}{r}$$

The force and the co-linear acceleration it produces are linked by Newton's second law of motion ($F = ma$), where the reduced mass m' is used:

Eq. 3.3 $$\frac{Ze^2}{r^2} = m'\frac{\mathrm{v}^2}{r} \qquad \text{or} \quad m'\mathrm{v}^2 = \frac{Ze^2}{r}$$

We can also obtain expressions for the energy of this two-body system. The kinetic energy is as expected:

Eq. 3.4 $$KE = \frac{1}{2}m'\mathrm{v}^2$$

The relevant potential energy is the electrostatic potential energy (as gravitational potential energy will be negligible):

Eq. 3.5 $$PE = \frac{Qq}{r} = -\frac{Ze^2}{r}$$

And the total energy of the system is the sum of these two:

* See Appendix: Constants

Eq. 3.6 $$E = KE + PE = \frac{1}{2} m' \mathrm{v}^2 - \frac{Ze^2}{r}$$

We now substitute Equation 3.3 into this expression:

Eq. 3.7 $$E = \frac{1}{2}\left(\frac{Ze^2}{r} \right) - \frac{Ze^2}{r} = -\frac{1}{2}\frac{Ze^2}{r}$$

This result shows that the total energy is a function of the orbital radius, with higher energy orbits being found at greater radii. However, there were indications from spectral analysis that only certain orbits might be available to the electron. There were also problems with a classical model involving an orbiting electron, which – because of its acceleration – must radiate energy. In such a system, the atom would constantly lose energy and the electron must spiral into the nucleus, destroying the atom.

Bohr got around this difficulty by postulating the **quantization of angular momentum** in the atom, requiring that the angular momentum of the orbiting electron (L) be an integral multiple of a constant: $h/2\pi \equiv \hbar$ (h-bar), where h is Planck's constant*. The quantization equation introduces the quantum number n:

Eq. 3.8 $L \equiv m' \mathrm{v}\, r = n\hbar$ where $n = 1, 2, 3, ...$

This requirement provides for discrete orbits in the atom, and an electron in one of these orbits does *not* radiate as predicted classically.

The de Broglie Wavelength

An alternative way to view this requirement was given in 1923 by Louis de Broglie, who proposed that matter could behave as a wave, just as waves sometimes behave as particles of matter. The **de Broglie wavelength** assigned to a particle is determined by the particle's momentum p, using the same relation that links the momentum and wavelength of a photon:

Eq. 3.9 $$p_{photon} = \frac{E}{c} = \frac{h}{\lambda}$$

Applying this relation to a particle of mass m' moving at a speed v gives the desired result:

Eq. 3.10 $$\lambda = \frac{h}{p} = \frac{h}{m'\mathrm{v}}$$

Now suppose the orbiting electron is exhibiting its wavelike properties; instead of a particle, we can imagine a wave pattern traveling around and around the orbit. At each point in the orbit, the amplitude of the wave will be the sum of all the waves that overlap there. If the phase of each overlapping wave is the same at a given point, then *constructive* interference will occur and a standing wave – equivalent to an existing particle – will be produced, as shown in Figure 3.2. However, if the phase of each successive wave is even slightly different, then the waves will

* See Appendix: Constants

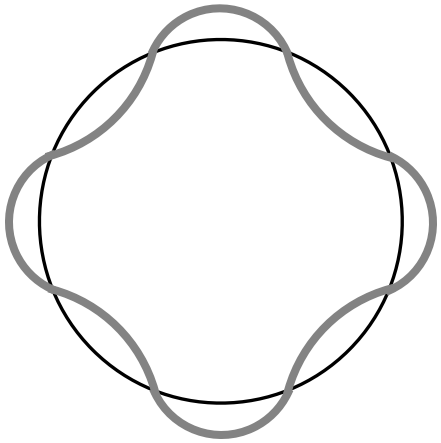
interfere *destructively* and no standing wave will occur; this is equivalent to saying that no particle can exist in such an orbit.

The condition for constructive interference is that the circumference of the orbit should be an integral number of de Broglie wavelengths:

Eq. 3.11 $$2\pi r = n\lambda = \frac{nh}{m'\mathrm{v}}$$

This leads immediately to the quantization requirement of Equation 3.8. Figure 3.2 shows the fourth orbit of the Bohr atom, for which the circumference is equal to four de Broglie wavelengths.

Figure 3.2: Standing waves in an orbit with n = 4



Now utilizing Equation 3.8 we can write the speed in terms of the radius:

Eq. 3.12 $$\mathrm{v} = \frac{n\hbar}{m'r}$$

Substituting this expression into Equation 3.3 gives the following:

Eq. 3.13 $$m'\left(\frac{n\hbar}{m'r}\right)^2 = \frac{Ze^2}{r}$$

Simplification of this equation yields the quantization of the orbital radii:

Eq. 3.14 $$r_n = \frac{n^2\hbar^2}{Ze^2m'}$$

Substitution of this expression into Equation 3.7 gives the quantization of the energy:

Eq. 3.15 $$E = -\frac{1}{2}\frac{Ze^2}{r} = -\frac{1}{2}Ze^2\left(\frac{Ze^2m'}{n^2\hbar^2}\right) = -\frac{Z^2e^4m'}{2\hbar^2n^2} = E_n$$

Thus both the orbital radius and the energy of one-electron atoms are quantized. We can simplify these rules as follows:

Eq. 3.16 $r_n = \frac{n^2}{Z} r_o$ where $r_o = \frac{\hbar^2}{e^2 m'}$

Eq. 3.17 $E_n = -\frac{Z^2}{n^2} E_o$ where $E_o = \frac{e^4 m'}{2\hbar^2}$

For hydrogen, $Z = 1$, giving the following values for the first few orbits:

$$r = r_o, 4r_o, 9r_o, 16r_o, \ldots \text{ and } E = -E_o, -\frac{E_o}{4}, -\frac{E_o}{9}, -\frac{E_o}{16}, \ldots$$

Figure 3.3 depicts the first four hydrogen orbits in the Bohr model, with relative radii of 1, 4, 9, and 16.

Figure 3.3: The first four hydrogen Bohr orbits (nucleus not shown)

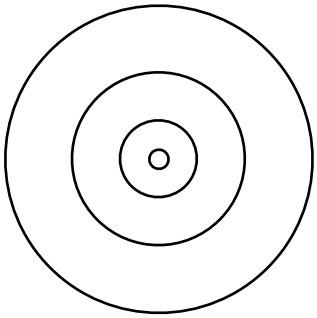
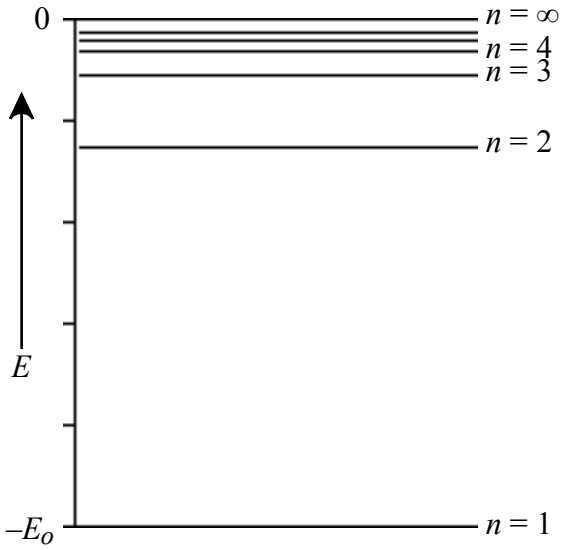


Figure 3.4 is a one-dimensional graph of the first six hydrogen energy levels predicted by the Bohr model; the 'points' on the graph are extended into horizontal lines, which will be utilized shortly.

Figure 3.4: Hydrogen energy levels



Note that *all* of the energy levels are negative; this is because each energy level represents a *bound* orbit, for which the total energy must be negative. Note also that as n increases to ∞, the

energy approaches zero; in practice, n never gets this big, as the radius – which increases as n^2 – would become infinitely large and the atom would fill the universe.

The same pattern of energy levels exists for other one-electron atoms (ions); the energy scale will be different however, being modified by different factors of Z and m'.

Electronic Transitions

An electron existing in one of these discrete orbits maintains a constant energy; however, the electron can change its energy by moving to a different orbit. A transition to a lower orbit will require that the electron lose energy, while transition to a higher orbit will involve an increase in energy. We can use Equation 3.17 to write an expression for the change in energy associated with a transition between two levels: the upper level (n) and the lower level (m). (Note that in this case, m is a quantum number, rather than mass.)

Eq. 3.18 $$\Delta E_{nm} = E_n - E_m = -Z^2 E_o \left(\frac{1}{n^2} - \frac{1}{m^2} \right)$$

For the hydrogen atom, $Z = 1$, and the energy change is as follows:

Eq. 3.19 $$\Delta E_{nm} = E_o \left(\frac{1}{m^2} - \frac{1}{n^2} \right)$$

As written, this quantity is always positive (because $n > m$). As noted above, this amount of energy must be *lost* in a downward electron transition or *gained* in an upward transition. The atom can gain or lose energy during collisions with other atoms or during interactions with photons. It is the photon interaction that interests us.

An atom may gain energy from the radiation field by *absorbing* a photon, or it may lose energy to the radiation field by *emitting* a photon. Whether performing absorption or emission, an electron undergoing a transition between two orbits must interact with a photon of a very specific energy – equal to the transition energy given by Equation 3.19.

Eq. 3.20 $$\Delta E_{nm} = E_{photon} = h\nu = hc/\lambda$$

This allows us to link the transition quantum numbers (n and m) to a particular photon wavelength:

Eq. 3.21 $$\frac{hc}{\lambda_{nm}} = Z^2 E_o \left(\frac{1}{m^2} - \frac{1}{n^2} \right)$$

We can solve this for the transition wavelength:

Eq. 3.22 $$\lambda_{nm} = \frac{hc}{Z^2 E_o} \left(\frac{1}{m^2} - \frac{1}{n^2} \right)^{-1} = \frac{2hc\hbar^2}{Z^2 e^4 m'} \left(\frac{1}{m^2} - \frac{1}{n^2} \right)^{-1} = \frac{4\pi\hbar^3 c}{Z^2 e^4 m'} \left(\frac{1}{m^2} - \frac{1}{n^2} \right)^{-1}$$

This formula can be simplified if we define the **Rydberg constant for hydrogen** (for which $Z = 1$):

Eq. 3.23 $$R_H \equiv \frac{e^4 m'}{4\pi \hbar^3 c}$$

This result produces the **Rydberg formula for hydrogen** (which was derived empirically from spectral data in 1888 – 25 years before the Bohr model):

Eq. 3.24 $$\lambda_{nm} = \frac{1}{R_H}\left(\frac{1}{m^2} - \frac{1}{n^2}\right)^{-1} \text{ or } \frac{1}{\lambda_{nm}} = R_H\left(\frac{1}{m^2} - \frac{1}{n^2}\right)$$

At this point, we will illustrate our progress by examining the numerical values of some of the quantities of interest for hydrogen. These, of course, depend on the values of the basic constants employed. Three sources are presented in Table 3.1 for comparison: Allen (1973), Cox (2000), and NIST (2006)*. Figures in *italics* are *calculated* values.

Table 3.1: Calculated values for hydrogen

Source	Allen 1973	Cox 2000	NIST 2006	
e	4.80325e–10	*4.8032068e–10*	*4.80320427e–10*	esu
eV	1.602192e–12	1.60217733e–12	1.602176487e–12	ergs
h	1.05459e–27	1.05457266e–27	1.054571628e–27	erg-s
c	29979250000	29979245800	29979245800	cm/s
m_p	1.672661e–24	1.6726231e–24	1.672621637e–24	g
m_n	___	1.6749286e–24	1.674927211e–24	g
m_e	9.10956e–28	9.1093897e–28	9.10938215e–28	g
m'	*9.10460e–28*	*9.1044313e–28*	*9.10442373e–28*	g
r_o	*5.29463e–9*	*5.2946545e–9*	*5.29465407e–9*	cm
E_o	*2.17874e–11*	*2.1786876e–11*	*2.17868542e–11*	ergs
E_o	*13.5985*	*13.59829*	*13.5982861*	eV
E_1	*–13.5985*	*–13.59829*	*–13.5982861*	eV
E_2	*–3.3996*	*–3.39957*	*–3.3995715*	eV
E_3	*–1.5109*	*–1.51092*	*–1.5109207*	eV
R_H	*109678*	*109677.58*	*109677.5835*	cm^{-1}
$1/R_H$	*911.759*	*911.76334*	*911.763341*	Å

As the precision with which the various constants are known continues to increase, the values of calculated quantities will change accordingly. We will generally not concern ourselves with attaining extremely precise results, but we will use enough precision to illustrate each concept sufficiently.

Now armed with the Rydberg formula and the Bohr model, we are prepared to investigate the various transitions available to the hydrogen electron. For historical reasons, these transitions

* See Appendix: Constants

have been grouped together into *series*, according to their lower levels (m). Those for which the lower level is the first level ($m = 1$) make up the Lyman series; $m = 2$ for the Balmer series, 3 for the Paschen series, 4 for the Brackett series, 5 for the Pfund series, and so on. For each series, the quantum number of the upper level may take on values ranging from $n = m + 1$ to $n = \infty$, which represents the **series limit**.

Figure 3.5: Hydrogen transitions – the Lyman and Balmer series

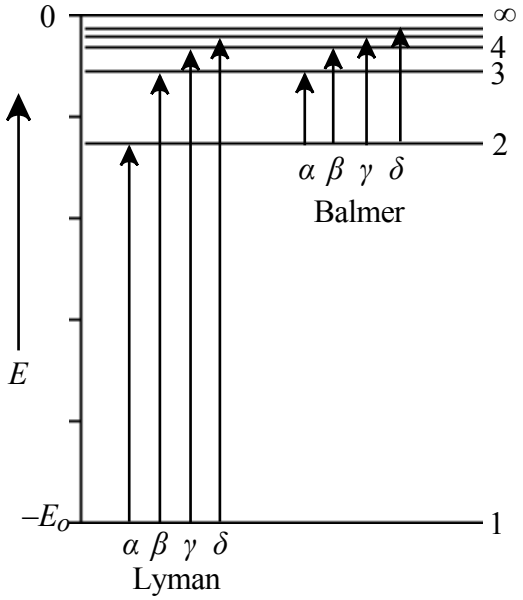


Figure 3.5 shows the first four transitions of the Lyman series (Lyα: 1→2, Lyβ: 1→3, Lyγ: 1→4, and Lyδ: 1→5) and the Balmer series (Hα: 2→3, Hβ: 2→4, Hγ: 2→5, and Hδ: 2→6). Wavelengths for each series can be calculated with the Rydberg formula (Equation 3.24). For the Balmer series, $m = 2$ and the relevant equation is as follows:

Eq. 3.25 $$\lambda_n = \frac{1}{R_H}\left(\frac{1}{4} - \frac{1}{n^2}\right)^{-1}$$

Using constants from each of the three sources previously noted, the Balmer wavelengths are calculated in Table 3.2.

We note that as the precision with which the various constants are known has improved over the years, the precision with which the Balmer wavelengths can be calculated has also improved.

We now compare the wavelengths from Table 3.2 with those found in essentially *any* book – such as Cox (2000) page 70, shown in the last column of Table 3.2. These values are all conspicuously *lower* than the calculated wavelengths in Table 3.2; one would expect better agreement, especially given the simplicity of the Rydberg formula. What could be wrong?

The answer turns out to be relatively straightforward. The values obtained by calculation with the Rydberg formula are *vacuum* wavelengths, while those listed in most books are the wavelengths in *air*. In spectroscopy, it is customary to report *air* wavelengths for $\lambda > 2000$ Å simply because most measurements are performed in air; however, *vacuum* wavelengths are usually given for $\lambda < 2000$ Å because photons in this region tend to be absorbed by air.

Table 3.2: Balmer wavelengths

Source	Allen (1973)	Cox (2000)	NIST (2006)		Cox (2000) p70	
R_H	109678	109677.58	109677.5835	cm^{-1}		
$1/R_H$	911.759	911.76334	911.763341	Å		
	calculated	*calculated*	*calculated*			tabulated
n	λ_n	λ_n	λ_n			λ_n
3	*6564.66*	*6564.6960*	*6564.69606*	Å	Hα	6562.80
4	*4862.71*	*4862.7378*	*4862.73782*	Å	Hβ	4861.32
5	*4341.71*	*4341.7302*	*4341.73020*	Å	Hγ	4340.46
6	*4102.92*	*4102.9350*	*4102.93503*	Å	Hδ	4101.73
∞	*3647.04*	*3647.0534*	*3647.05337*	Å	Limit	3646

The conversion from the calculated *vacuum* wavelength to the measured *air* wavelength involves the **index of refraction** of air – usually denoted n (not to be confused with the quantum number n). Because the value of the index of refraction varies with temperature and pressure, we normally use the value for air at *standard temperature and pressure*: 15°C and 1 atmosphere. Cox (2000) gives several formulae for this index, but one has incorrect coefficients. The formula we will use is as follows:

Eq. 3.26 $$n = 1 + A + \frac{B}{C - \sigma^2} + \frac{D}{E - \sigma^2} \text{ where } \sigma \equiv 1/\lambda$$

The coefficient values depend on the units of λ as shown in Table 3.3.

Table 3.3: Index of refraction coefficients

Cox (2000) page:	69	170 (modified)	262
lambda units:	cm	Å	nm
A	6.43280e–5	6.43280e–5	6.43280e–5
B	2.94981e+6	2.94981e–10	2.94981e–8
C	1.46000e+10	1.46000e–6	1.46000e–4
D	2.55400e+4	2.55400e–12	2.55400e–10
E	4.10000e+9	4.10000e–7	4.10000e–5

Table 3.4: Balmer wavelengths in air (calculated with NIST (2006) values)

	λ (vacuum)	n	λ (air)	Moore (1972)	difference
Hα	6564.69606	1.00027624	6562.88312	6562.817	0.066
Hβ	4862.73782	1.00027934	4861.37984	4861.332	0.048
Hγ	4341.73020	1.00028114	4340.50990	4340.468	0.042
Hδ	4102.93503	1.00028222	4101.77743	4101.737	0.040
limit	3647.05337	1.00028497	3646.01436	3645.981	0.033

Now, applying the modified coefficients (for λ in Å) to determine the value of the index of refraction at each of the Balmer wavelengths, we can obtain the air wavelengths shown in Table 3.4. These are compared with the observed Balmer wavelengths given in Moore (1972) and found to be quite reasonable.

There are a few related terms that can be mentioned here (with values from NIST (2006)*):

- The **Bohr radius (a_o)** is the radius of the first orbit of an infinite mass H atom (for which the reduced mass is equal to the electron mass): $a_o = \dfrac{\hbar^2}{e^2 m_e} = 0.5291772083$ Å
- The **Rydberg constant for an infinite mass atom (R_∞)** is calculated by setting the reduced mass equal to the electron mass: $R_\infty \equiv \dfrac{e^4 m_e}{4\pi\hbar^3 c} = 109737.31568549\ \text{cm}^{-1}$

Other units of energy – besides ergs and eVs – occasionally are utilized:

- The **Rydberg (ryd)**: $hcR_\infty = 2.17987190 \times 10^{-11}$ erg $= 13.60569172$ eV
- The **atomic unit:** $e^2/a_o = 2hcR_\infty = 2$ ryd $= 27.21138344$ eV

We are interested in the absorption of photons by matter. Which photons a hydrogen atom absorbs will depend on what level the electron occupies. Given a set of hydrogen atoms, how many of them have an electron in level 1, level 2, level 3, etc.? The answer to this question depends on the **statistical weight** of each level – the number of different ways an electron may occupy the level – which is obtained from quantum numbers. We can find the appropriate quantum numbers by examining a **quantum mechanical model** of the hydrogen atom, obtained by solving the **Schrödinger equation**.

The Quantum Mechanical Model

Quantum mechanics was developed to provide a means of understanding the physics of very small objects – such as atoms. We will make use of the particle/wave duality of matter and presume that the state of the hydrogen atom can be described by a **wave function**, which will be the solution to a differential equation that is based on the energy of the atom. To begin, let us note that the total energy of the atom is equal to the sum of the kinetic energy and the potential energy (as in Equation 3.6):

Eq. 3.27 $KE + PE = E$

We now substitute expressions for the kinetic and potential energies:

Eq. 3.28 $PE = V(\vec{r},t)$ (Potential energy is a function of position vector and time.)

Eq. 3.29 $KE = \dfrac{1}{2} m' \vec{v}^2 = \dfrac{\vec{p}^2}{2m'}$ ($\vec{p}$ is the momentum vector.)

Inserting these into Equation 3.27 produces a classical expression for the energy (in which the left side of the equation is known as the **Hamiltonian**):

---

* See Appendix: Constants

Eq. 3.30 $$\frac{\vec{p}^2}{2m'} + V(\vec{r},t) = E$$

Now we replace the momentum ($\vec{p}$) and energy (E) with their equivalent quantum mechanical **operators**. (Operators perform some operation – such as differentiation – on the wave function to yield the quantity in question.) These operators are as follows:

- $\vec{p} \Rightarrow -i\hbar\nabla$
- $\vec{p}^2 \Rightarrow -\hbar^2\nabla^2$ (where ∇^2 is the *Laplacian*)
- $E \Rightarrow i\hbar\frac{\partial}{\partial t}$

These will operate on the **wave function**, $\Psi(\vec{r},t)$. Inserting these into Equation 3.30 transforms it into the **Schrödinger equation**:

Eq. 3.31 $$-\frac{\hbar^2}{2m'}\nabla^2\Psi(\vec{r},t) + V(\vec{r},t)\Psi(\vec{r},t) = i\hbar\frac{\partial\Psi(\vec{r},t)}{\partial t}$$

This form of the equation is perfectly general, allowing for potential functions of any type. However, very often we will deal with a potential function that is time-independent – meaning that $V = V(\vec{r})$. If this is the case, then *energy eigenfunctions* exist; this means that the operator – in this case the Hamiltonian – operating on the wave function Ψ will produce a constant – in this case, the total energy E – multiplied by the wave function ($H\Psi = E\Psi$, where Ψ is the **eigenfunction** and E is the **eigenvalue**). These eigenfunctions will have the following form:

Eq. 3.32 $$\Psi(\vec{r},t) = u(\vec{r})e^{-iEt/\hbar}$$

We can illustrate this by inserting this wave function and the time-independent potential into Equation 3.31 and simplifying:

Eq. 3.33 $$-\frac{\hbar^2}{2m'}\nabla^2 u(\vec{r})e^{-iEt/\hbar} + V(\vec{r})u(\vec{r})e^{-iEt/\hbar} = i\hbar\frac{\partial u(\vec{r})e^{-iEt/\hbar}}{\partial t}$$

Eq. 3.34 $$-\frac{\hbar^2}{2m'}e^{-iEt/\hbar}\nabla^2 u(\vec{r}) + V(\vec{r})u(\vec{r})e^{-iEt/\hbar} = i\hbar u(\vec{r})\frac{de^{-iEt/\hbar}}{dt} = i\hbar u(\vec{r})\left(\frac{-iE}{\hbar}\right)e^{-iEt/\hbar}$$

Dividing through by the exponential eliminates time from the equation:

Eq. 3.35 $$-\frac{\hbar^2}{2m'}\nabla^2 u(\vec{r}) + V(\vec{r})u(\vec{r}) = E\,u(\vec{r})$$ or

Eq. 3.36 $$\nabla^2 u(\vec{r}) + \frac{2m'}{\hbar^2}\left[E - V(\vec{r})\right]u(\vec{r}) = 0$$

This is the **time-independent Schrödinger equation**, a three-dimensional differential equation. Its solution is $u(\vec{r})$, the spatial component of the wave function, which is a function of the three spatial coordinates. We are free to choose a suitable three-dimensional coordinate

system to use in describing the hydrogen atom; most people find spherical polar coordinates (*r, θ, ϕ*) to be the most logical choice.

Having selected a coordinate system, we can now write the Laplacian; for spherical polar coordinates it is as follows (applied to our function $u(\vec{r})$):

Eq. 3.37 $$\nabla^2 u(\vec{r}) = \frac{1}{r}\frac{\partial^2}{\partial r^2}[r\,u(\vec{r})] + \frac{1}{r^2}\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right]u(\vec{r})$$

We will now define the angular portion of the Laplacian (inside the brackets [...]) as the *operator* Ω (Note: this Ω is *not* a solid angle). Equation 3.37 is then simplified:

Eq. 3.38 $$\nabla^2 u(\vec{r}) = \frac{1}{r}\frac{\partial^2}{\partial r^2}[r\,u(\vec{r})] + \frac{1}{r^2}\Omega u(\vec{r})$$

The Schrödinger equation is then written as follows:

Eq. 3.39 $$\frac{1}{r}\frac{\partial^2}{\partial r^2}[r\,u(\vec{r})] + \frac{1}{r^2}\Omega u(\vec{r}) + \frac{2m'}{\hbar^2}\left[E - V(\vec{r})\right]u(\vec{r}) = 0$$

We now make two assumptions about our problem: first, that the potential function is spherically symmetric ($V(\vec{r}) = V(r)$), having no angular dependence; and second, that we can write the spatial portion of the wave function as the product of a radial function (*R*) and an angular function (*Y*), in hopes of performing **separation of variables**:

Eq. 3.40 $$u(\vec{r}) = R(r)Y(\theta,\phi)$$

Making these substitutions into Equation 3.39 yields the following:

Eq. 3.41 $$\frac{1}{r}\frac{\partial^2}{\partial r^2}[r\,R(r)Y(\theta,\phi)] + \frac{1}{r^2}\Omega R(r)Y(\theta,\phi) + \frac{2m'}{\hbar^2}\left[E - V(r)\right]R(r)Y(\theta,\phi) = 0$$

The partial derivatives operate only on their respective components of $u(\vec{r})$:

Eq. 3.42 $$Y(\theta,\phi)\frac{1}{r}\frac{\partial^2}{\partial r^2}[r\,R(r)] + \frac{R(r)}{r^2}\Omega\,Y(\theta,\phi) + \frac{2m'}{\hbar^2}\left[E - V(r)\right]R(r)Y(\theta,\phi) = 0$$

Next we divide through the equation by $R(r)Y(\theta, \phi)/r^2$:

Eq. 3.43 $$r^2\left\{\frac{1}{rR(r)}\frac{\partial^2}{\partial r^2}[r\,R(r)] + \frac{2m'}{\hbar^2}\left[E - V(r)\right]\right\} + \frac{\Omega\,Y(\theta,\phi)}{Y(\theta,\phi)} = 0$$

We now note that each of the two terms in this equation contains different variables; the first term has only the radial coordinate while the second term has only angular coordinates. If we move the angular term to the right side of the equation, it will be clear that the left side is dependent only on *r* and its derivatives while the right side is dependent only on *θ* and *ϕ* and their derivatives.

Eq. 3.44 $$r^2\left\{\frac{1}{rR(r)}\frac{\partial^2}{\partial r^2}[r\,R(r)]+\frac{2m'}{\hbar^2}\big[E-V(r)\big]\right\}=-\frac{\Omega\,Y(\theta,\phi)}{Y(\theta,\phi)} \quad \text{or} \quad f(r)=g(\theta,\phi)$$

These two completely independent functions can only be equal for all values of r, θ, and ϕ if each function is equal to the same *constant*. With incredible foresight, we will set this constant equal to $\ell(\ell+1)$. This gives us two separate equations – one radial and one angular:

Eq. 3.45 $$r^2\left\{\frac{1}{rR(r)}\frac{\partial^2}{\partial r^2}[r\,R(r)]+\frac{2m'}{\hbar^2}\big[E-V(r)\big]\right\}=\ell(\ell+1) \quad \text{(the radial equation)}$$

Eq. 3.46 $$-\frac{\Omega\,Y(\theta,\phi)}{Y(\theta,\phi)}=\ell(\ell+1) \qquad \text{(the angular equation)}$$

We will rewrite the radial equation, changing the partial derivatives into full derivatives:

Eq. 3.47 $$\frac{1}{r}\frac{d^2}{dr^2}[r\,R(r)]+\frac{2m'}{\hbar^2}\big[E-V(r)\big]R(r)=\ell(\ell+1)\frac{R(r)}{r^2}$$

The angular equation requires further separation of variables. We begin by assuming that the angular component of the wave function can be written as a product of two angular functions, each dependent only on a single variable:

Eq. 3.48 $$Y(\theta,\,\phi)=\Theta(\theta)\,\Phi(\phi)$$

Inserting this into Equation 3.46, and also using the full form of Ω, we find the following:

Eq. 3.49 $$\frac{-1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}[\Theta(\theta)\Phi(\phi)]\right)-\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}[\Theta(\theta)\Phi(\phi)]=\ell(\ell+1)\Theta(\theta)\Phi(\phi)$$

We can again separate variables by multiplying through by $\dfrac{\sin^2\theta}{\Theta(\theta)\Phi(\phi)}$ and rearranging, collecting each variable on opposite sides of the equation:

Eq. 3.50 $$-\frac{\sin\theta}{\Theta(\theta)}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\Theta(\theta)\right)-\ell(\ell+1)\sin^2\theta=\frac{1}{\Phi(\phi)}\frac{\partial^2}{\partial\phi^2}\Phi(\phi)$$

Now the left side is a function of θ and its derivatives, while the right side is a function of ϕ and its derivatives. This can only be true if each side equals the same constant, one that we will call $-m_\ell^{\,2}$. Inserting this constant and converting partial derivatives gives two angular equations:

Eq. 3.51 $$\frac{d^2}{d\phi^2}\Phi(\phi)=-m_\ell^{\,2}\Phi(\phi)$$

Eq. 3.52 $$-\frac{1}{\sin\theta}\frac{d}{d\theta}\left(\sin\theta\frac{d}{d\theta}\Theta(\theta)\right)+\frac{m_\ell^2\Theta(\theta)}{\sin^2\theta}=\ell(\ell+1)\Theta(\theta)$$

These two equations, together with Equation 3.47 can be solved independently for the individual components of the wave function, which can then be multiplied together to give the complete wave function.

We begin with Equation 3.51 because it is the simplest and can be solved by inspection:

Eq. 3.53 $\Phi(\phi) = C_1 e^{im_\ell(\phi + C_2)}$

We may choose the integration constant C_2 to be 0, as there are no special ϕ values in the atom. C_1 is a normalization constant, to be determined shortly.

Eq. 3.54 $\Phi(\phi) = C_1 e^{im_\ell \phi} = C_1(\cos m_\ell \phi + i \sin m_\ell \phi)$

We will now put restrictions on the value of m_ℓ by applying the boundary condition that $\Phi(\phi) = \Phi(\phi + 2\pi)$; that is, the wave function must be in phase with itself after wrapping around the atom (recall the de Broglie requirement that the circumference of an orbit be an integral number of wavelengths). This can be illustrated by examining the real part of the function:

Eq. 3.55 $\cos(m_\ell \phi) = \cos(m_\ell(\phi + 2\pi)) = \cos(m_\ell \phi + 2\pi m_\ell)$

This condition will be met if $m_\ell = 0, \pm1, \pm2, \pm3$, etc. – in short, if m_ℓ is an integer. We can then identify m_ℓ as a quantum number associated with the coordinate ϕ. This component of the wave function can then be written as follows:

Eq. 3.56 $$\Phi_{m_\ell}(\phi) = \frac{1}{\sqrt{2\pi}} e^{i m_\ell \phi} = \frac{1}{\sqrt{2\pi}} (\cos m_\ell \phi + i \sin m_\ell \phi)$$

The value of the normalization constant $\left(C_1 = 1/\sqrt{2\pi}\right)$ is found by requiring that the probability of the electron existing *somewhere* be equal to 1. Mathematically, this is done by setting $\int_0^{2\pi} \Phi^* \Phi \, d\phi = 1$ (where Φ^* signifies the complex conjugate of Φ) and solving for C_1.

The solution for $\Theta(\theta)$ is not so simple. It can be found to be the following:

Eq. 3.57 $$\Theta_{\ell m_\ell}(\theta) = \left[\frac{(\ell - m)!}{(\ell + m)!} \frac{2\ell + 1}{2}\right]^{1/2} P_\ell^m(\cos\theta)$$

Here $P_\ell^m(\cos \theta)$ is the Associated Legendre Function; and ℓ must be an integer, such that it takes on the following values:

Eq. 3.58 $\ell = |m_\ell|, |m_\ell| + 1, |m_\ell| + 2, \ldots$

Only for these values does the wave function remain finite. (Infinite wave functions are highly impractical.)

Legendre Function

$$P_\ell(x) = \frac{1}{2^\ell \ell!}\frac{d^\ell}{dx^\ell}\left(x^2-1\right)^\ell \qquad \text{(Spiegel 1968, p146)}$$

Associated Legendre Function

$$P_\ell^{\,m}(x) = \left(1-x^2\right)^{m/2}\frac{d^m P_\ell(x)}{dx^m} \qquad \text{(Spiegel 1968, p149)}$$

The solution of the radial equation requires that a potential energy function be defined; the most appropriate function in this case is the electrostatic potential from Equation 3.5:

Eq. 3.59 $$V(r) = -\frac{Ze^2}{r}$$

Insertion of this expression into Equation 3.47 and solving yields the following:

Eq. 3.60 $$R_{n\ell}(r) = -\left(\frac{2Z}{na_o}\right)^{3/2}\left\{\frac{(n-\ell-1)!}{2n\left[(n+\ell)!\right]^3}\right\}^{1/2} e^{-\frac{Zr}{na_o}}\left(\frac{2Zr}{na_o}\right)^\ell L_{n+\ell}^{2\ell+1}\left(\frac{2Zr}{na_o}\right)$$

The function $L_{n+\ell}^{2\ell+1}\left(\frac{2Zr}{na_o}\right)$ is an Associated Laguerre Polynomial, with the allowed values of n being $n = \ell+1, \ell+2, \ell+3, \ell+4$ ... .

Laguerre Polynomial

$$L_n(x) = e^x\frac{d^n}{dx^n}\left(x^n e^{-x}\right) \quad \text{where } n = 0, 1, 2, \ldots \qquad \text{(Spiegel 1968, p153)}$$

Associated Laguerre Polynomial

$$L_n^m(x) = \frac{d^m}{dx^m}L_n(x) \quad \text{where } n, m = 0, 1, 2, \ldots \qquad \text{(Spiegel 1968, p155)}$$

The angular components of the wave function combine to give what are known as the **spherical harmonics** $Y_{\ell m_\ell}(\theta, \phi)$:

Eq. 3.61 $$Y_{\ell m_\ell}(\theta, \phi) = \Theta_{\ell m_\ell}(\theta)\,\Phi_{m_\ell}(\phi)$$

The spherical harmonics are independent of the potential ($V(r)$). They are eigenfunctions of the operator L^2 and also of L_z, where $\vec{L}$ is the orbital angular momentum operator $\left(\vec{L} = \vec{r}\times\vec{p}\right)$:

Eq. 3.62 $$L^2 Y_{\ell m_\ell}(\theta, \phi) = \ell(\ell+1)Y_{\ell m_\ell}(\theta, \phi) \qquad \text{and}$$

Eq. 3.63 $$L_z Y_{\ell m_\ell}(\theta, \phi) = m_\ell\, Y_{\ell m_\ell}(\theta, \phi)$$

We can then write the complete wave function for hydrogen:

Eq. 3.64 $\Psi_{n\ell m_\ell}(r, \theta, \phi) = R_{n\ell}(r)\, Y_{\ell m_\ell}(\theta, \phi)$

The three quantum numbers take on the following values:

- $n = 1, 2, 3, ...$
- $\ell = 0, 1, 2, ...(n-1)$
- $m_\ell = 0, \pm1, \pm2, ... \pm \ell$

Examples of some of these wave functions are given by Rybicki & Lightman (1979):

Eq. 3.65 $$\Psi_{100} = \frac{1}{\sqrt{\pi}}\left(\frac{Z}{a_o}\right)^{3/2} e^{-Zr/a_o}$$ (for $n = 1$, $\ell = 0$, and $m_\ell = 0$)

Eq. 3.66 $$\Psi_{200} = \frac{1}{4\sqrt{2\pi}}\left(\frac{Z}{a_o}\right)^{3/2}\left(2 - \frac{Zr}{a_o}\right) e^{-Zr/2a_o}$$ (for $n = 2$, $\ell = 0$, and $m_\ell = 0$)

Eq. 3.67 $$\Psi_{32\pm1} = \frac{1}{81\sqrt{\pi}}\left(\frac{Z}{a_o}\right)^{3/2}\left(\frac{Zr}{a_o}\right)^2 e^{-Zr/3a_o} \sin\theta\cos\theta\, e^{\pm i\phi}$$ (for $n = 3$, $\ell = 2$, and $m_\ell = \pm1$)

Quantum Numbers

Each quantum number is associated with a different spatial coordinate:

- $n \rightarrow r$ n is the **principal quantum number**, giving the total energy.
- $\ell \rightarrow \theta$ ℓ is the **azimuthal quantum number**, giving the orbital angular momentum.
- $m_\ell \rightarrow \phi$ m_ℓ is the **magnetic quantum number**.

For each value of n, there are n values of ℓ.

For each value of ℓ, there are $2\ell+1$ values of m_ℓ .

For each value of n, there are n^2 degenerate eigenfunctions (having the same energy), if the energy depends only on the value of n.

Quantum number combinations for the first three orbits are shown in Table 3.5. Note that for $n = 1$, there is only one combination; for $n = 2$, there are four combinations; for $n = 3$, there are nine combinations. This means that the **statistical weight** of each **level** n – the number of different ways for an electron to be in that level – is n^2. Almost.

Table 3.5: Examples of valid quantum number combinations:

n	ℓ	m_ℓ
1	0	0
2	0	0
2	1	0, ±1
3	0	0
3	1	0, ±1
3	2	0, ±1, ±2

There is one more quantum number. Electrons have a property called **spin**. The spin quantum number is labeled s, and its value is *always* the same: $s = {}^1/_2$. Now spin is just a name; the electrons are not really spinning. In fact, spin has no classical analog – it is purely quantum mechanical. However, the manner in which spin enters into our discussion of the atom is somewhat reminiscent of a classical situation in astronomy.

A planet orbiting the Sun clearly has orbital angular momentum. If the planet is also rotating, it will have some rotational angular momentum as well. These two quantities can be added together vectorially to produce the total angular momentum for the planet.

In the atom, we already have the orbital angular momentum of the electron, linked to the quantum number ℓ. We now add another component (spin) that also contributes to the angular momentum; we will refer to this contribution as the **spin angular momentum**. Thus, the quantum numbers ℓ and s play somewhat similar roles in our discussion; and just as ℓ is linked to the quantum number m_ℓ, the spin s will be linked to another quantum number called m_s.

We have seen that m_ℓ may take on values up to $\pm\ell$ and that m_ℓ values differ by integers; the same is true for m_s, which may take on values up to $\pm s$, with m_s values differing by integer amounts. Because s always has a value of ${}^1/_2$, the only allowable values of m_s are $\pm {}^1/_2$. These are sometimes denoted as **spin up** and **spin down**; they are also referred to as **polarization states**.

Specification of the **state** of a hydrogen atom thus requires values for four quantum numbers: n, ℓ, m_ℓ, and m_s. We have previously found that the statistical weight of a particular energy level n is n^2, based on the first three quantum numbers. With the inclusion of spin, we find that for each combination of spatial quantum numbers – for each electron *orbital* – there are two values for m_s, doubling the overall statistical weight ($g_n = 2n^2$).

With some understanding of the hydrogen atom, we are now ready to step into the real world to consider the electronic structure of multi-electron atoms.

Multi-electron Atoms

Multi-electron atoms are more complicated, not only because there are more electrons to track, but also because the electrons interact with each other, through Coulomb's law. Fortunately we are able to construct a successful model of such an atom by building on our existing quantum numbers and the structure they represent.

Pauli Exclusion Principle

In a multi-electron atom, each electron has a set of four quantum numbers that define its position and momentum. According to the **Pauli exclusion principle**, no two electrons may have the same set of quantum numbers; this is equivalent to requiring that no two electrons may be in the same place, doing the same thing, at the same time.

The quantum numbers already described are n, ℓ, m_ℓ, and m_s; but they may also be defined in a slightly different manner, as n, ℓ, j, and m_j. These relate to each other as follows:

Eq. 3.68 $\quad j = \ell \pm s$

Eq. 3.69 $m_j = m_\ell + m_s$

With this set, the quantum number j represents a combination of the *orbital* angular momentum and the *spin* angular momentum – the two quantities most critical to defining the electron's state. This set provides a different – but equivalent – way of looking at the atom.

Total Quantum Numbers

In a multi-electron atom, we need not keep track of each individual electron's angular momentum because the important quantities are the *totals* for all electrons in the atom.

Each electron has an orbital angular momentum vector $\vec{\ell}$ (characterized by the quantum number ℓ) and a spin angular momentum vector $\vec{s}$ (characterized by the quantum number s). In a multi-electron atom, these vectors add (vectorially) to produce a total angular momentum for the atom. There are two different ways of achieving this total.

We could vectorially add all of the individual orbital angular momentum vectors to obtain a **resultant orbital angular momentum vector** $\vec{L} = \sum \vec{\ell}$, and we could do the same for the spin to find the **resultant spin vector** $\vec{S} = \sum \vec{s}$. Then we could add these vectors to obtain the **total angular momentum vector** $\vec{J} = \vec{L} + \vec{S}$. This is called **LS coupling** (or **Russell-Saunders coupling**), and it is valid for elements with low Z.

For higher Z elements, we would first find a combined angular momentum for each electron ($\vec{j} = \vec{\ell} + \vec{s}$) and then add these to obtain the total angular momentum $\vec{J} = \sum \vec{j}$. This is called **jj coupling**. (Our work in this book will normally involve LS coupling.)

We now introduce the quantum numbers L, S, and J associated with the vectors $\vec{L}, \vec{S}$, and $\vec{J}$. These must be distinguished from their individual-electron counterparts ℓ, s, and j.

- ℓ represents the orbital angular momentum of a *single* electron in the atom.
- L represents the *total* orbital angular momentum of *all* the electrons in the atom.
- Electrons with $\ell = 0$ are called s electrons. (Note: This s is not spin.)
- Electrons with $\ell = 1$ are called p electrons.
- Electrons with $\ell = 2$ are called d electrons.
- Electrons with $\ell = 3$ are called f electrons.

These letters refer to *sharp*, *principal*, *diffuse*, and *fundamental* – terms derived from the appearance of spectral lines. They can be applied to the figures from Table 3.5 to describe the types of electrons each **shell** – designated by the principal quantum number n and labeled by still another letter – can hold.

• The K ($n = 1$) shell can hold 2 s electrons	total = 2 electrons
• The L ($n = 2$) shell can hold 2 s and 6 p electrons	total = 8 electrons
• The M ($n = 3$) shell can hold 2 s, 6 p, and 10 d electrons	total = 18 electrons
• The N ($n = 4$) shell can hold 2 s, 6 p, 10 d, and 14 f electrons	total = 32 electrons
	total = $2n^2$ electrons

Electron Configurations

In general, an atom will not have the appropriate number of electrons to fill each of its shells. The **ground state** of an atom – the lowest possible energy configuration – will normally involve filled inner shells with an unfilled outer shell. We can specify what type of electrons make up the ground state of an element by listing the **electron configuration** as $n\ell^{\#}$, where # is the number of electrons with the n and ℓ values given. Examples are given in Table 3.6.

Table 3.6: Ground state configurations

Element	Configuration
H	$1s^1$
He	$1s^2$
Li	$1s^2\ 2s^1$
Be	$1s^2\ 2s^2$
B	$1s^2\ 2s^2\ 2p^1$

The key point in analyzing multi-electron atoms is that although the angular momenta of individual electrons change continually, the *total* angular momentum is constant. (The individual ℓ values may change, but the value of L remains the same.) How do the ℓ values combine to give L? How do the s values combine to give S? We can find the answer by examining the values of $M_L = \sum m_\ell$ and $M_S = \sum m_s$. The process is as follows:

From the electron configuration, we have the values of n and ℓ; the latter gives the allowable values of m_ℓ. These can be added to give the value of M_L, and this quantity determines the value of L for the atom. A similar analysis of the spin will yield the values of M_S and S.

For example, consider the ground state of hydrogen, which has only one electron: H $1s^1$. From the electron configuration we have $n = 1$, and the single s electron has $\ell = 0$. The only value of m_ℓ available for $\ell = 0$ is $m_\ell = 0$. With only one electron, the value of M_L will be the same as m_ℓ (= 0) and the minimum L value required for this M_L is also 0.

Spin is treated in a similar manner, but the numbers are usually different. With only one electron, s can only be $^1/_2$. This means that m_s may be either $+\,^1/_2$ or $-\,^1/_2$. The value of M_S will then be either $\pm\,^1/_2$, and the minimum value of S required to produce this M_S value is $^1/_2$.

In summary, for the ground state of hydrogen, we have $L = 0$ and $S = {}^1/_2$.

The ground state of helium is slightly more complicated because there are now two electrons. From the electron configuration ($1s^2$) we again have $n = 1$, and each s electron has $\ell = 0$. The only value of m_ℓ available for $\ell = 0$ is $m_\ell = 0$, and this is true for both electrons. The resulting value of M_L will again be 0, and the minimum L value required for this M_L is also 0.

Spin is quite different this time. For each electron, s can only be $^1/_2$, and once again m_s may be either $+\,^1/_2$ or $-\,^1/_2$. But because these two electrons already have the *same* values for n (= 1), ℓ (= 0), and m_ℓ (= 0), they cannot have the same value of m_s (by the Pauli exclusion principle). As there are only two different values for m_s, and there are two electrons, clearly one must have

spin up ($m_s = +{}^1/_2$) and the other must have spin down ($m_s = -{}^1/_2$). The makes $M_S = 0$, and the minimum value of S required to produce this M_S value is 0. (Note that if the two electrons had the same spin, then $S = 1$ would be possible, but this situation is prohibited by the Pauli exclusion principle.)

Therefore, for the ground state of helium, we have $L = 0$ and $S = 0$, a result slightly different from that for hydrogen.

We are now in a position to provide a designation for the electron configuration that will indicate the manner in which these electrons combine to produce orbital angular momentum and spin angular momentum. This designation, which is based on the values of L and S, is called a **term**. It has the general form ^{2S+1}L, where L is a *letter* (*S, P, D, F, G, H*, ...) corresponding to the *numerical value* of L (0, 1, 2, 3, 4, 5, ...), and the quantity $2S+1$ is the **multiplicity**.

For the ground state of hydrogen, the multiplicity is $2({}^1/_2) + 1 = 2$, making the term designation ^2S (pronounced *doublet S*). For the ground state of helium, the multiplicity is 2(0) + 1, and the term designation is ^1S (*singlet S*). (Note: The S in ^2S and ^1S has nothing to do with spin. The spin information is included in the multiplicity, where we will find triplets, quartets, quintets, sextets, etc. in addition to singlets and doublets.)

Equivalent and Non-equivalent Electrons

The above treatment for helium considered its two electrons to be **equivalent**, meaning they share the same values for n and ℓ. **Non-equivalent electrons** – those that differ in either n or ℓ – must be handled differently. An example can be found in the excited state of helium. Suppose that one of the two electrons is excited to the second level ($n = 2$), giving an electron configuration of He $1s^1\ 2s^1$. What term applies to this configuration?

Such problems are most easily solved by constructing a table showing the possible values of m_ℓ and m_s, along with the resulting values of M_L and M_S, the required values of L and S, and the terms that these values combine to produce. Table 3.7 shows a solution to the current problem.

Table 3.7: Terms for two non-equivalent s electrons

electron:	1*s*		2*s*				
state	m_ℓ	m_s	m_ℓ	m_s	M_L	M_S	group
1	0	$+{}^1/_2$	0	$+{}^1/_2$	0	+1	#1
2	0	$+{}^1/_2$	0	$-{}^1/_2$	0	0	#1
3	0	$-{}^1/_2$	0	$-{}^1/_2$	0	–1	#1
4	0	$-{}^1/_2$	0	$+{}^1/_2$	0	0	#2

Because each electron is an s electron, the only possible value of m_ℓ is 0, making $M_L = 0$ for each combination. Because the electrons are non-equivalent, all four possible combinations of spins are allowed, giving four different **states**. (A **state** is a combination in which all four quantum numbers are specified: n, ℓ, m_ℓ, m_s. The states are numbered in the first column of this table, but this is only for reference.) These states produce the different values for M_S shown.

We now determine what values of L and S are implied by these values for M_L and M_S. We begin by noting the largest value of M_S (= 1). This will require a minimum value of S = 1, resulting in a multiplicity of 2(1) + 1 = 3 , indicating a triplet term. To find the corresponding value of L, we locate the largest value of M_L in a state with M_S = +1; as there is only one such state in this simple example, this is an easy task: M_L = 0. The minimum value of L needed to produce a state with M_L = 0 is 0, giving us a ^3S term.

Each term may be composed of multiple states. In this case, the ^3S term contains three states, each with the same value of M_L (0) and a different value of M_S (–1, 0, +1). These states are identified in Table 3.7 as group #1, consisting of states 1, 2, and 3, but we could just as well have made our ^3S term using states 1, 3, and 4. It really does not matter as states 2 and 4 have the same values for M_L and M_S, and that is all that concerns us. Our first result then is that two non-equivalent electrons produce a ^3S term.

But there is more. We accounted for three of the four states with our ^3S term, but there is one state left over. It has M_L = 0, which leads to another S term, and it has M_S = 0, which gives a multiplicity of 2(0) + 1 = 1, resulting in a ^1S term. Such a term consists of only one state (that is what *singlet* means), one with M_L = 0 and M_S = 0, which is just what we have in state 4. Thus we identify group #2 as the state(s) required for a ^1S term.

While the only term for two equivalent s electrons (designated s^2) is ^1S, the terms for two non-equivalent s electrons (designated ss) are ^1S and ^3S. This result is typical; equivalent electrons (s^2, p^4, d^3, etc.) will always produce fewer terms than their corresponding non-equivalent electrons (ss, $pppp$, ddd, etc.). Also note that there are non-equivalent electron combinations (such as sp, sd, spd, etc.) that have no equivalent counterparts.

We will now consider a more complex example: pp (two non-equivalent p electrons). For each electron, ℓ = 1, giving m_ℓ = 0, ±1; and $s = {}^1/_2$, giving $m_s = \pm {}^1/_2$ – signified by ↑ (spin up) or ↓ (spin down). We consider all possible combinations in Table 3.8, as we are not limited by the Pauli exclusion principle; this yields 36 states. (Note: The *order* in which the states are listed does not matter, but it can make the bookkeeping easier.) As before, we calculate M_L and M_S and then assign states to terms.

We begin by finding the highest M_L , which is 2. This requires a term with L = 2, which would be a D term. The highest M_S that goes with M_L = 2 is M_S = 1, giving a multiplicity of 2(1) + 1 = 3. This means we need a ^3D term, which will consist of 3×5 = 15 states: 5 for the 5 different M_L values (0, ±1, ±2) available for L = 2, and 3 for the 3 different M_S values (0, ±1) available for S =1. (In general, this product is $(2L+1)(2S+1)$, and it is equal to the *statistical weight* of the *term*.) A suitable set of 15 states is identified as group **#1**.

Next we find the highest M_L of the remaining states; this is again 2, which will give us another D term. The highest M_S with M_L = 2 is M_S = 0, which means we will have S = 0 and a multiplicity of 1, leading to a ^1D term. Such a term requires 1×5 = 5 states; a suitable set is group **#2**.

Table 3.8: Terms for two non-equivalent p electrons (pp)

$m_{\ell 1}$			$m_{\ell 2}$					
+1	0	–1	+1	0	–1	M_L	M_S	group
↑			↑			2	1	#1
↑			↓			2	0	#1
↓			↓			2	–1	#1
↓			↑			2	0	#2
↑				↑		1	1	#1
↑				↓		1	0	#1
↓				↓		1	–1	#1
↓				↑		1	0	#2
↑					↑	0	1	#1
↑					↓	0	0	#1
↓					↓	0	–1	#1
↓					↑	0	0	#2
	↑		↑			1	1	#3
	↑		↓			1	0	#3
	↓		↓			1	–1	#3
	↓		↑			1	0	#4
	↑			↑		0	1	#3
	↑			↓		0	0	#3
	↓			↓		0	–1	#3
	↓			↑		0	0	#4
	↑				↑	–1	1	#1
	↑				↓	–1	0	#1
	↓				↓	–1	–1	#1
	↓				↑	–1	0	#2
		↑	↑			0	1	#5
		↑	↓			0	0	#5
		↓	↓			0	–1	#5
		↓	↑			0	0	#6
		↑		↑		–1	1	#3
		↑		↓		–1	0	#3
		↓		↓		–1	–1	#3
		↓		↑		–1	0	#4
		↑			↑	–2	1	#1
		↑			↓	–2	0	#1
		↓			↓	–2	–1	#1
		↓			↑	–2	0	#2

Of the remaining states, the highest M_L is 1, with a maximum M_S of 1, giving a ^3P term. Its $3 \times 3 = 9$ states are labeled as group **#3**.

Of the remaining states, the highest M_L is 1, with a maximum M_S of 0, giving a ^1P term. Its $1 \times 3 = 3$ states are labeled as group **#4**.

Of the remaining states, the highest M_L is 0, with a maximum M_S of 1, giving a ^3S term. Its $3 \times 1 = 3$ states are labeled as group **#5**.

The only remaining state has $M_L = 0$ and $M_S = 0$, giving a ^1S term. This $1 \times 1 = 1$ last state is labeled as group **#6**.

So the terms for *pp* electrons are ^3D, ^1D, ^3P, ^1P, ^3S, and ^1S.

For comparison, we will now determine the terms for p^2 electrons; here we must obey the Pauli exclusion principle and only allow states in which the two electrons have different m_ℓ *or* different m_s. Table 3.9 lists the 15 allowed states.

Table 3.9: Terms for two equivalent p electrons (p^2)

m_ℓ					
+1	0	–1	M_L	M_S	group
↑↓			2	0	#1
↑	↑		1	1	#2
↑	↓		1	0	#1
↑		↑	0	1	#2
↑		↓	0	0	#1
↓	↑		1	0	#2
↓	↓		1	–1	#2
↓		↑	0	0	#2
↓		↓	0	–1	#2
	↑↓		0	0	#3
	↑	↑	–1	1	#2
	↑	↓	–1	0	#1
	↓	↑	–1	0	#2
	↓	↓	–1	–1	#2
		↑↓	–2	0	#1

The analysis proceeds as before. Of all the states, the highest M_L is 2, with a maximum M_S of 0, giving a ^1D term. Its $1 \times 5 = 5$ states are labeled as group **#1**.

Of the remaining states, the highest M_L is 1, with a maximum M_S of 1, giving a ^3P term. Its $3 \times 3 = 9$ states are labeled as group **#2**.

The only remaining state has $M_L = 0$ and $M_S = 0$, giving a ^1S term. This $1 \times 1 = 1$ last state is labeled as group **#3**.

So the terms for p^2 electrons are ^1D, ^3P, and ^1S.

As a general rule, singlet terms include $M_S = 0$

doublet terms include $M_S = +{}^1/_2, -{}^1/_2$

triplet terms include $M_S = +1, 0, -1$

quartet terms include $M_S = +{}^3/_2, +{}^1/_2, -{}^1/_2, -{}^3/_2$ etc.

an S term includes $M_L = 0$

a P term includes $M_L = +1, 0, -1$

a D term includes $M_L = +2, +1, 0, -1, -2$ etc.

When faced with a set of terms resulting from a particular electron configuration, we often want to know which of them produces the lowest energy. For terms given by *equivalent* electrons, **Hund's rule** applies: The terms with greatest multiplicity will have the lowest energy, and of those, the terms with the greatest L will have the lowest energy. This says that for the example of p^2 electrons given above, the ${}^3P$ term will have the lowest energy.

There is more to the story. Recall that $\vec{J} = \vec{L} + \vec{S}$. The vectors $\vec{L}$ and $\vec{S}$ will add such that the vector $\vec{J}$ has integer length (if S is an integer) or half-integer length (if S is half-integer). In general, J has the following values:

Eq. 3.70 $$J = (L+S),(L+S-1),(L+S-2),\ldots|L-S| \quad \text{for } J \geq 0$$

For $L = 0$ and $S = {}^1/_2$, $J = {}^1/_2$ $2S + 1 = 2$

For $L = 1$ and $S = {}^1/_2$, $J = {}^3/_2, {}^1/_2$ $2S + 1 = 2$

For $L = 3$ and $S = {}^3/_2$, $J = {}^9/_2, {}^7/_2, {}^5/_2, {}^3/_2$ $2S + 1 = 4$

For $L = 2$ and $S = 1$, $J = 3, 2, 1$ $2S + 1 = 3$

Note: While L and S specify a term (${}^{2S+1}L$), L, S, and J specify a **level** (${}^{2S+1}L_J$).

For ground state hydrogen, we had an electron configuration of $1s^1$, which produced a ${}^2S$ term, for which $L = 0$ and $S = 1/2$. Adding vectors with these lengths can only produce a $\vec{J}$ vector of length $J = {}^1/_2$, resulting in a level designation of ${}^2S_{1/2}$. (Note: This is still called a doublet even though only one J value occurs. The number of J values is $2S+1$ if $L > S$ and $2L+1$ if $L < S$.)

For ground state helium, we had an electron configuration of $1s^2$, which produced a ${}^1S$ term, for which $L = 0$ and $S = 0$. Adding vectors with these lengths can only produce a $\vec{J}$ vector of length $J = 0$, resulting in a level designation of ${}^1S_0$.

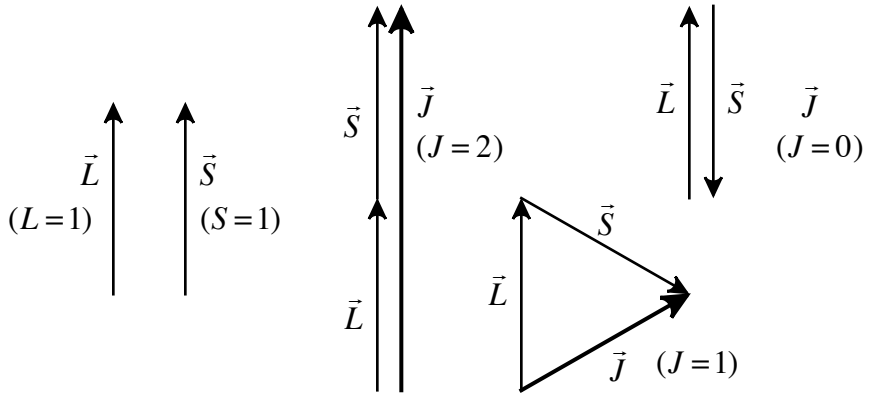
All closed (filled) shells have ${}^1S_0$ ground states, for which $L = 0$ and $S = 0$. For example, consider neon ($Z = 10$), which has a ground state configuration of $1s^2\ 2s^2\ 2p^6$. Table 3.10 lists all possible states – there is only one – for the two $1s$ electrons, the two $2s$ electrons, and the six $2p$ electrons. Because the shell is filled, each spin-up electron is paired with a spin-down electron, insuring that $M_S = 0$. Similarly, for every electron with $m_\ell = +1$, there is another with $m_\ell = -1$, insuring that $M_L = 0$.

Table 3.10: Ground state for neon (Z = 10)

	1*s*	2*s*	2*p*						
m_ℓ:	0	0	+1	0	–1	M_L	M_S	term	level
	↑↓	↑↓	↑↓	↑↓	↑↓	0	0	^1S	$^1S_0$

As another example, consider a ^3P term, for which $L = 1$ and $S = 1$. In this case, J can take on three different integer values (2, 1, 0), as shown in Figure 3.6. This leads to the three different levels that make up the triplet: $^3P_2$, $^3P_1$, and $^3P_0$.

Figure 3.6: Determination of J values for a ^3P term



Which energy level J is the lowest? The rule is simple: If the shell is *less* than half filled, the *lowest J level* will have the lowest energy; if the shell is *more* than half filled, then the *highest J level* will have the lowest energy. As examples, consider the ground states of carbon and oxygen, as shown in Figure 3.7.

Figure 3.7: Ground states for carbon (regular) and oxygen (inverted)

Carbon (Z = 6): $1s^2\ 2s^2\ 2p^2 \Rightarrow {}^3P_0$ Oxygen (Z = 8): $1s^2\ 2s^2\ 2p^4 \Rightarrow {}^3P_2$

The fourth quantum number was called the magnetic quantum number. This is because the application of an external magnetic field will split each J level into states with different M_J, where $M_J = J, J-1, J-2, \ldots -J$, and the statistical weight of a J level is then $g_J = 2J + 1$.

Where does all of this lead? How does it involve absorption and emission?

Absorption and emission involve electron transitions between different energy states within the atom. A reasonably precise vocabulary has been developed to insure that we can say what we mean as we discuss these processes:

- A **state** is totally specified by four quantum numbers, either [L, S, J, M_J] or [L, S, M_L, M_S] . An example of a state is $^2P_{3/2}$ $M_J = -1/2$.
- A **level** specifies only the quantum numbers [L, S, J] and thus may include several states. An example of a level is $^2P_{3/2}$.
- A **term** specifies only the quantum numbers [L, S] and thus may include several levels. An example of a term is ^2P.
- Transitions between **states** are called **components**.
- Transitions between **levels** are called **lines**.
- Transitions between **terms** are called **multiplets**.

Charlotte Moore's *A Multiplet Table of Astrophysical Interest* (1972) lists transitions for each element and ionization stage, grouped by multiplet. For example, consider the second multiplet for neutral magnesium.

Mg I multiplet #2 is given as 3 $^3P^\circ \rightarrow$ 4 ^3S (the number preceding each term indicates the value of n). The three lines that make up this multiplet are as follows:

transition	wavelength
$3\ ^3P_2^\circ \rightarrow 4\ ^3S_1$	5183.6042
$3\ ^3P_1^\circ \rightarrow 4\ ^3S_1$	5172.6843
$3\ ^3P_0^\circ \rightarrow 4\ ^3S_1$	5167.3216

Each of these transitions has the same upper level (4 $^3S_1$), but the lower levels are different. As this is a regular multiplet, the lower level with the lowest energy is found in the third transition (3 $^3P_0^\circ$), which has the shortest wavelength and therefore the largest energy jump to the upper level.

Application of a magnetic field would split these lines into components.

Note that there is an extra designation for each of the lower levels in this multiplet; this is for **parity**, which may be either **odd** ($^3P^\circ$) or **even** (^3P). Parity refers to the symmetry of the wave function with respect to the origin. Parity is said to be even if $\Psi(-x) = \Psi(x)$, and parity is odd if $\Psi(-x) = \Psi(-x)$.

For one electron, the parity P (not to be confused with the $L = 1$ term) is given by $P \approx (-1)^\ell$; for even values of ℓ (even parity), $P \approx +1$, while for odd values of ℓ (odd parity), $P \approx -1$. For example, s and d electrons have even parity while p and f electrons have odd parity.

For a multi-electron atom, the parity is given by $P \approx (-1)^{\Sigma\ell}$, where the exponent is the *algebraic sum* of the individual ℓ values. The parity of a term depends on the *configuration* from which it arises, rather than on the term itself. A ^3P term may be either odd or even: an sp configuration (odd) can produce a $^3P^\circ$ term while a pp configuration (even) can produce ^3P.

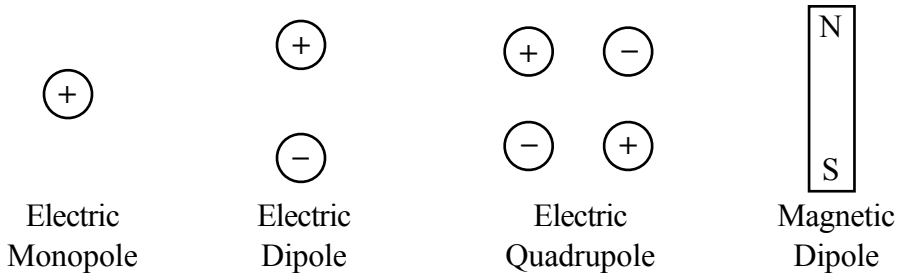
Even configurations include *s, d, ss, sd, pp, ssd*, etc.

Odd configurations include *p, sp, pd, ppp, spd*, etc.

Radiative Transitions

We now have notation for identifying different terms and levels, and we are ready to examine transitions of the electron from one level to another, as will occur during absorption and emission. Although there are numerous energy levels within most atoms, the electronic transitions that can occur are not unrestricted. Quantum mechanics provides us with **selection rules**, which govern the changes in the quantum numbers that can occur during a radiative transition. There are different types of transitions, and each type has a different set of selection rules. The most common type of transition is the **electric dipole** transition; other, less common, types include **magnetic dipole** transitions and **electric quadrupole** transitions. Figure 3.8 shows the differences in the various charge configurations.

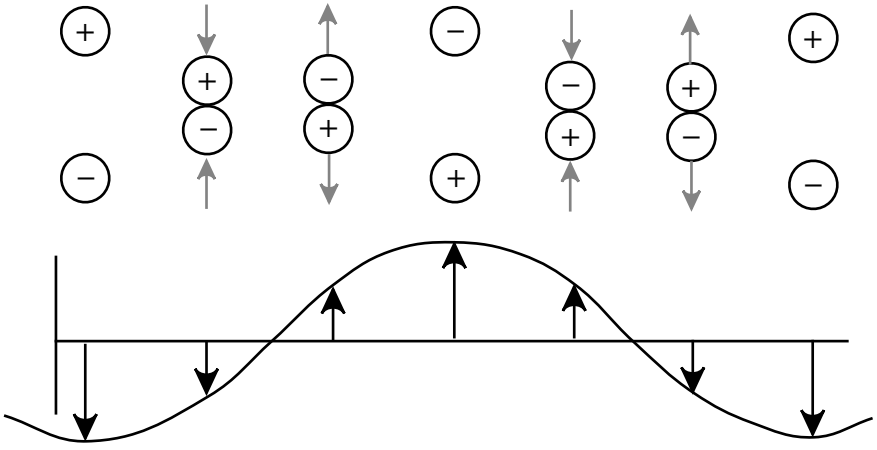
Figure 3.8: Multipole configurations



Radiation is produced by oscillations of the different charge configurations. Monopoles do not oscillate, but oscillation of the charges in the electric dipole (as shown in Figure 3.9) will produce an oscillation in the direction and strength of the electric field vector surrounding the dipole, which will generate sinusoidal electromagnetic waves that propagate outwards from the dipole.

Similarly, magnetic dipole and electric quadrupole oscillations can also generate electromagnetic waves, but the details are different from (and more complex than) those for the electric dipole.

Figure 3.9: An oscillating electric dipole and its changing electric field



Selection rules arise from quantum mechanics. Recall that the probability of the electron existing is given by $\int \Psi^* \Psi d\vec{r}$, where Ψ is the wave function, and $d\vec{r}$ signifies integration over

all space. Electrons in different energy levels have different wave functions; the probability of an electron making a transition from an initial state Ψ_i to a final state Ψ_f is given by $\int \Psi_f^* \vec{d}\, \Psi_i \, d\vec{r}$, where the operator $\vec{d}$ depends on the type of transition. If the value of this integral is zero, then the transition cannot proceed by the method associated with the given operator. We will not attempt to derive selection rules here, but will simply present the results for electric dipole transitions.

Electric dipole selection rules:

1	$\Delta J = 0, \pm 1$	but $J = 0$ cannot combine with $J = 0$
2	$\Delta M = 0, \pm 1$	
3	Parity change	
4	$\Delta \ell = \pm 1$	(one electron jumps)

and for strict LS coupling,

5	$\Delta S = 0$	(multiplicity is constant)
6	$\Delta L = 0, \pm 1$	but $L = 0$ cannot combine with $L = 0$

"**Allowed transitions**" follow these selection rules. Transitions that violate rules 1 through 4 are called "**forbidden transitions**". This does not mean that they cannot happen, but rather that they occur less frequently, by other mechanisms – usually either magnetic dipole or electric quadrupole transitions. We will find later that the probabilities for magnetic dipole or electric quadrupole transitions are orders of magnitude lower than for electric dipole transitions.

To illustrate the electric dipole selection rules, consider which levels can combine with $^2P_{3/2}$, for which $S = {}^1/_2$, $L = 1$, and $J = {}^3/_2$.

By rule #5, all terms will be doublets, and by rule #6 they must have $L = 0$, 1, or 2. As the given term has even parity, only odd parity terms will be able to combine with it. This means we need consider only $^2P°$, $^2S°$, and $^2D°$ terms.

By rule #1, J must be ${}^1/_2$, ${}^3/_2$, or ${}^5/_2$. We now apply these to the above terms to determine the allowable levels.

For $^2S°$, $L = 0$ and $S = {}^1/_2$. There is only one way to combine these, and the resulting J value is ${}^1/_2$. Therefore, the only level for this term is $^2S°_{1/2}$.

For $^2P°$, $L = 1$ and $S = {}^1/_2$. Now L and S may add to make ${}^3/_2$ or subtract to make ${}^1/_2$. This produces two levels for the $^2P°$ term: $^2P°_{3/2}$ and $^2P°_{1/2}$.

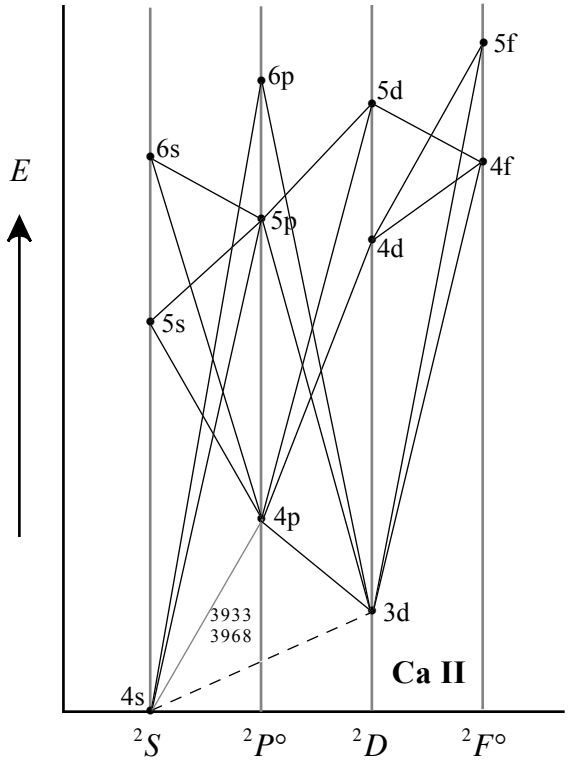
For $^2D°$, $L = 2$ and $S = {}^1/_2$. L and S may add to make ${}^5/_2$ or subtract to make ${}^3/_2$. This produces two levels for the $^2D°$ term: $^2D°_{5/2}$ and $^2D°_{3/2}$.

A useful way of diagraming the transitions for a given element is the **Grotrian diagram** – a plot of energy vs. term. Information that may be included on such a diagram includes electron configurations, terms, energy levels, transition wavelengths, multiplet numbers, etc. Examples of Grotrian diagrams can be found in *Lang's Handbook* (1980) and in Herzberg (1944).

Ionized calcium's single electron results in a fairly simple diagram (Figure 3.10). In the ground state, the first three shells are completely filled, with only a single electron in the fourth shell (4*s*); the terms it forms are all doublets. Absorption transitions can excite this outer

electron to higher energy levels, but the transitions must obey the appropriate selection rules – normally those for electric dipole transitions. The ground state does not connect directly to the 5*s* or 6*s* levels, nor to those with *d* or *f* configurations, due to selection rule #6. The one apparent exception to this rule – and also to rule #3 – is the 4*s*–3*d* transition, which is a forbidden transition – indicated by a dashed line.

Figure 3.10: Grotrian diagram for Ca II



With two electrons in its outer shell, neutral calcium presents a considerably more complicated diagram (Figure 3.11). The first point to note is that Ca I forms singlets and triplets, which are segregated on the diagram. Normally, the transitions obey this segregation of terms as prescribed by selection rule #5, but there are a few exceptions – indicated by dashed lines.

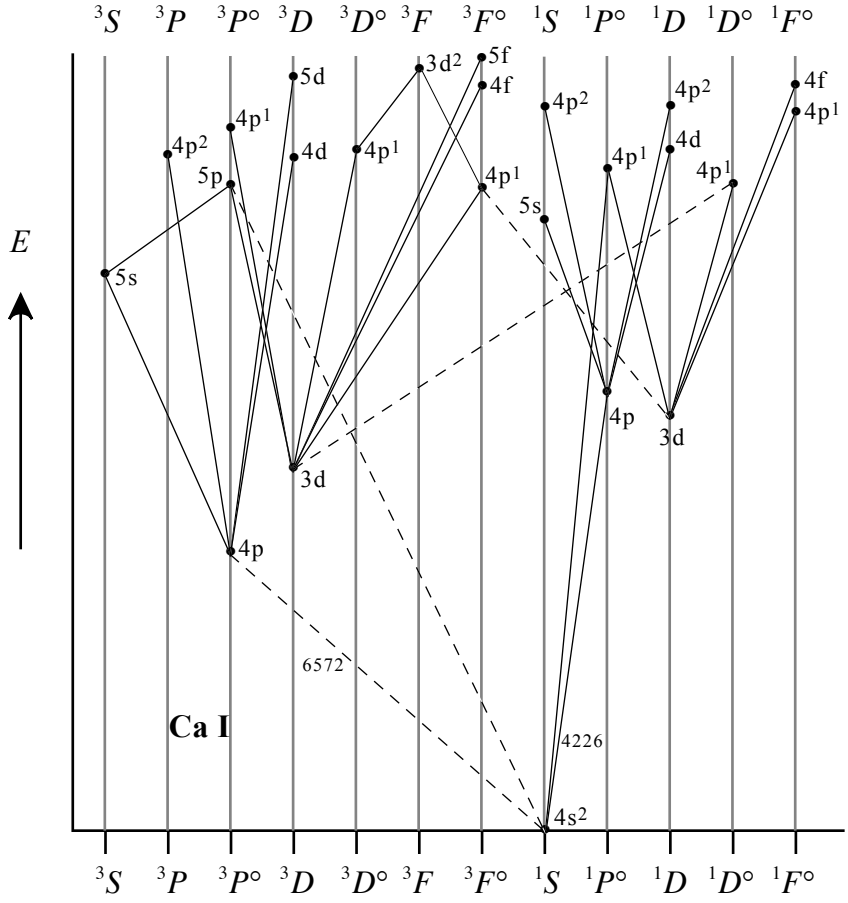
The ground state configuration is $4s^2$, meaning both electrons are *s* electrons. This level connects to levels marked 4*p*, for which the configuration is 4*s*4*p*. These levels in turn connect to 5*s* (4*s*5*s*), 4*d* (4*s*4*d*), 5*d* (4*s*5*d*), and $4p^2$ ($4p^2$) – for which *both s* electrons have been excited. The 3*d* levels do not connect to the ground state because changing one electron from *d* to *s* would violate selection rule #6.

The levels marked $4p^1$ each connect to a 3*d* level, implying that the configuration for these is 3*d*4*p* – different from the 4*p* levels (4*s*4*p*).

The terms along the horizontal axis normally alternate in parity. This makes the diagram neater because parity must change for electric dipole transitions with strict LS coupling.

Numbers on a few of the transitions indicate the principal wavelengths of a multiplet. Different *J* levels are not distinguished on the diagram, so each transition represents a multiplet.

Figure 3.11: Grotrian diagram for Ca I



In summary, we have reasonable models of simple atoms and a means for understanding the electronic structure of multi-electron atoms, along with terminology for describing electronic transitions. The next task is to determine just how the electrons will be distributed within the atom under any given conditions. We know the possible places they *might* be, but we do not yet know where they *are*. This will be the subject of the next chapter.

CHAPTER 4: Thermodynamic Equilibrium

In this chapter, we will try to determine how electrons distribute themselves within the atom. We know what configurations are *possible* for the electrons to attain, but we have not discussed which ones they are *most likely* to attain. We can imagine that the electron configurations of individual atoms will somehow be related to the energy of the gas as a whole, and this energy can be characterized by the average kinetic energy of a particle of the gas. The parameter we use to measure this average kinetic energy is the **temperature**, or more precisely, the **kinetic temperature** of the gas. The relation is relatively simple:

Eq. 4.1 $\overline{KE} = \overline{\left(\frac{1}{2} m \mathrm{v}^2\right)} = \frac{3}{2} kT$ ergs/particle (k is Boltzmann's constant*)

We can then link this temperature to the distribution of electrons within the atom, but only under certain conditions. Because electronic transitions occur so rapidly, the atoms are constantly changing their configurations. However, given sufficient time, the atoms will reach a state in which no *net* transitions occur. This state will be called **thermodynamic equilibrium**, and it will be characterized by the temperature of the gas. The standard assumption is that without any external input of energy or work, a system will move toward a state of thermodynamic equilibrium.

There are several different processes by which a gas may change its state, and we may establish equilibrium distribution functions that will describe how the particles of a gas in equilibrium at some temperature T will arrange themselves. The first of these processes is excitation.

Excitation

We have already mentioned the **ground state**, which is the configuration of electrons in the atom that has the lowest possible energy. **Excitation** refers to the population of higher electronic energy levels in the atom; **excited states** of an atom have greater energy than the ground state. Generally there are many excited states in an atom, but only one ground state. Obviously, in order for an atom to make a transition from the ground state to an excited state, it will have to obtain additional energy from somewhere. The most common sources of such energy are **collisions**.

---

* See Appendix: Constants

Consider a gas of atoms at some kinetic temperature; this means that the atoms are moving, with an average kinetic energy determined by the temperature. As the atoms move in random directions, they collide with each other, exchanging energy in the process. The energy exchange may simply amount to a change in the *speeds* of the colliding particles, as is seen in the collisions of billiard balls. But the atom may also use some of the energy gained in a collision to *excite* its electrons to higher energy states, effectively storing some of this newly acquired energy. However, most gases are dense enough and their particle speeds are high enough that collisions are quite frequent; collisional energy stored an the excitation process may be swiftly lost in the next collision. Thus atoms are constantly being collisionally excited and de-excited, meaning that at any given time, the atoms in a gas will be distributed between the ground state and the various excited states.

The higher the temperature of the gas, the faster the particles will be moving; collisions that occur will be more frequent and more violent, and capable of exciting electrons to very high energy levels. We would then expect a hot gas to maintain a greater fraction of its atoms in excited states, with a greater fraction of those excited states requiring high energies to attain. Conversely, in a cool gas the collisional energies will be lower and less apt to result in excitation, leading to a greater fraction of the atoms being found in the ground state. We need a way to quantitatively describe the distribution of atoms among the various energy states at a given temperature.

Boltzmann Equation

For simplicity, consider a gas of hydrogen atoms. Let N_n be the number of atoms per cubic centimeter (the **number density**) with the electron in level n, and let N_m be the number density of atoms with the electron in level m. Let g_n and g_m be the statistical weights of levels n and m, and let E_n and E_m be the respective energies of these levels. Then the relative populations of these two groups of atoms are given by the **Boltzmann equation**:

Eq. 4.2
$$\frac{N_n}{N_m} = \frac{g_n}{g_m}\frac{e^{-E_n/kT}}{e^{-E_m/kT}} = \frac{g_n}{g_m}e^{-(E_n-E_m)/kT}$$

(Note that n and m need not be restricted to values of the principal quantum number; they may represent any two energy states.)

We are often interested in relating a particular excited state to the ground state. In this case, we will let m represent the ground state, and denote it by setting m equal to 0. (It is customary to denote the ground state with 0, letting g_0 be the statistical weight of the ground state, etc. However, the ground state of the hydrogen atom is $n = 1$; hence, we may regard g_1 as g_0 and N_1 as N_0 .) The Boltzmann equation relating the excited state n to the ground state is then as follows:

Eq. 4.3
$$\frac{N_n}{N_o} = \frac{g_n}{g_o}e^{-(E_n-E_o)/kT} = \frac{g_n}{g_o}e^{-\Delta E_n/kT}$$

The Boltzmann Hotel

We may provide an analogy for the Boltzmann equation in the form of the **Boltzmann Hotel**. First, let us write the essence of the Boltzmann equation: $N_n \approx g_n\, e^{-E_n/kT}$.

Our Boltzmann Hotel is a high-rise structure with a number of floors; suppose we are considering obtaining a room on the n^{th} floor.

We would like to know the number of occupants of the n^{th} floor (N_n). This will of course depend on the number of rooms on the n^{th} floor, multiplied by the number of persons each room is allowed to hold, which will give us the maximum capacity of the n^{th} floor (g_n). In general, the greater the capacity of a floor, the more occupants it is likely to have when the hotel is full, or nearly so.

However, it may be that the hotel is not full, that the floors are not necessarily near capacity. New guests at the hotel may choose to stay on *any* of the floors. Of course, room prices are not all equal; in general, rooms that are higher up are more expensive, due to the better view. The price of a room on the n^{th} floor is represented by E_n.

What else determines how guests will distribute themselves among the different floors? It would seem that money (T) would be a factor. If guests have little money, they will have to stay on one of the lower floors, but if they have plenty, they can stay wherever they want. Even better, let T represent the current state of the economy. If the economy is strong, there is plenty of money to be had, and at least some hotel guests will populate the upper levels of the hotels they visit. But if the economy is weak, then what guests the hotels do have will be staying near the ground floor because they cannot afford the higher rooms.

Partition Function

The Boltzmann equation relates two states to each other, but what we would really like to know is how a particular state relates to the *total* number density of atoms. We can obtain this total (N) by summing over all of the energy states n:

Eq. 4.4 $$N = \sum_n N_n = \frac{N_o}{g_o} \sum_n g_n e^{-\Delta E_n/kT}$$

The sum in this equation is known as the **partition function *U*(*T*)**:

Eq. 4.5 $$U(T) = \sum_n g_n e^{-\Delta E_n/kT}$$

The partition function is not easily obtained, as it involves a sum over all the energy levels in the atom, requiring knowledge of their energies and statistical weights. In fact, the sum diverges for finite values of T. However, as atoms are not infinitely large, we can usually truncate the sum after a reasonable number of terms (electrons are usually found relatively close to the nucleus).

Just as the statistical weight of a level gives the number of different combinations of quantum numbers that place an electron in that level, the partition function is essentially a statistical

weight for the whole atom, giving the number of different ways the atom can exist at a given temperature. It includes the ground state and all the excited states, weighted by Boltzmann equations.

We can now write an expression for the fraction of *all* hydrogen atoms that have their electrons in level *n*:

Eq. 4.6 $$\frac{N_n}{N} = \frac{g_n}{U(T)} e^{-\Delta E_n / kT}$$

Historical note: In the days before pocket calculators, astronomers relied primarily on common logs rather than natural logs. And of course in calculations involving atoms, energies are usually given in electron volts. This combination made it useful to define a new variable as follows:

$$\theta \equiv 1.602 \times 10^{-12} \left(\frac{ergs}{eV} \right) \frac{\log_{10} e}{kT} \approx \frac{5040}{T}$$

With this substitution, the exponential in Equation 4.6 is changed: $e^{-\Delta E_n / kT} \Rightarrow 10^{-\theta \Delta E_n}$, where E is in electron volts. This substitution is not essential now, but much of the literature still uses θ as a temperature variable. (Note: This θ has *nothing* to do with angles.)

As an example, we can calculate the partition function for a simple atom: Ca II (singly ionized calcium). We will need the energies and statistical weights of the different levels (identified by their configurations and terms), which can be obtained from Moore (1972) and from Grotrian diagrams from Lang (1980) (see Figure 3.10). (Recall from Chapter 3 that the statistical weight of a *term* is the product $(2S+1)(2L+1)$). Results are shown in Table 4.1.

Table 4.1: Terms, energies, and statistical weights for Calcium II

			$f(T) = g\, e^{-E/kT}$				
Term	E (eV)	g	f(2000)	f(3000)	f(4000)	f(5000)	f(6000)
$4s\ ^2S$	0	2	2	2	2	2	2
$3d\ ^2D$	1.69	10	5.51e–4	1.45e–2	7.42e–2	1.98e–1	3.81e–1
$4p\ ^2P°$	3.11	6	8.74e–8	3.58e–5	7.24e–4	4.40e–3	1.47e–2
$5s\ ^2S$	6.44	2	1.18e–16	3.04e–11	1.54e–8	6.45e–7	7.79e–6
$4d\ ^2D$	7.02	10	2.04e–17	1.61e–11	1.43e–8	8.40e–7	1.27e–5
$5p\ ^2P°$	7.48	6	8.50e–19	1.63e–12	2.26e–9	1.73e–7	3.13e–6
$4f\ ^2F°$	8.4	14	9.53e–21	1.08e–13	3.65e–10	4.78e–8	1.23e–6
$6s\ ^2S$	8.73	2	2.01e–22	4.32e–15	2.00e–11	3.17e–9	9.29e–8
$5d\ ^2D$	8.98	10	2.35e–22	8.21e–15	4.85e–11	8.88e–9	2.87e–7
$6p\ ^2P°$	9.20	6	3.94e–23	2.10e–15	1.54e–11	3.20e–9	1.12e–7
$5f\ ^2F°$	9.63	14	7.58e–24	9.30e–16	1.03e–11	2.75e–9	1.14e–7

Adding each column of $f(T)$ values yields the partition function for each temperature:

Table 4.2: The partition function (U(T)) and the fraction of atoms in the ground state (N_o/N)

T:	2000	3000	4000	5000	6000	7000	8000
U(*T*):	2.00055	2.01452	2.07497	2.20235	2.39527	2.64187	2.92838
N_o/N:	0.9997	0.9928	0.9639	0.9081	0.8350	0.7570	0.6830

It is immediately obvious that the partition function is dominated by the ground state at relatively low stellar temperatures – that is, $U(T) \approx g_0$. As the temperature goes up, the fraction of excited atoms increases as the partition function expands. In the atmospheres of cool stars, we could safely approximate $U(T)$ by g_0, but this approximation becomes less reliable at higher temperatures.

Of course, this is based on data for only one species – Ca II. While other atoms will exhibit similar behavior, the details will depend on the energies of the different levels – specifically how close to the ground state the excited states lie. Atoms with a large energy jump to the first excited state are more apt to be found in the ground state, even at relatively high temperatures. Atoms with low-lying excited states will be excited comparatively easily.

It should be remembered that the Boltzmann equation only predicts relative level populations for gases in *equilibrium*; however, this condition is usually – but not always – met in the atmospheres of stars.

Radiative Transition Probabilities

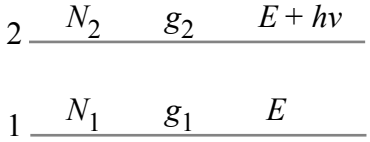
The Boltzmann equation describes the distribution of electrons over the various energy levels for a collection of atoms in thermodynamic equilibrium, but it says nothing about how the atoms arrive at that distribution. As already noted, they may achieve this distribution by the process of collision, and if the gas is sufficiently dense, the collision process will be very efficient. However, atoms may also undergo electronic transitions by absorbing or emitting suitable photons, and this brings up the question of the probability of radiative transitions.

We may divide radiative transitions into two groups:

- Spontaneous transitions, which occur without any external stimulus; and
- Induced transitions, which result from an interaction with the photon field.

Consider a two-level atom, as shown in Figure 4.1. The lower level (1) has energy E, statistical weight g_1, and population N_1 (atoms per cc), while the upper level (2) has energy $E + h\nu$, statistical weight g_2, and population N_2.

Figure 4.1: The two-level atom



Let P_{21} be the probability per second of a transition from level 2 to level 1, and let P_{12} be the probability per second of a transition from level 1 to level 2.

There is only one spontaneous transition to consider: an atom with its electron in level 2 may spontaneously emit a photon while making a transition to level 1. We may define a coefficient

A_{21} to be the probability per second for **spontaneous emission**, giving us $P_{21} = A_{21}$. A_{21} is known as the **Einstein A-coefficient**.

Absorption is obviously an induced process, as it requires the presence of photons having energy *hv* in order for the transition to proceed. Because photons are *necessary*, we expect the probability of an absorption transition to be proportional to the intensity at the specified frequency (I_v); this gives us $P_{12} = B_{12} I_v$, where B_{12} is the **Einstein B-coefficient**.

There is a third type of transition to consider. While absorption cannot occur *without* photons, it *is* possible for photons to stimulate the emission of additional photons of the same frequency, in the same direction, and in phase with the original photons; this is known as **stimulated emission**. Because this process is induced by the photon field, its probability will have the same dependence on the intensity as absorption did. We will write $P_{21} = B_{21} I_v$, where B_{21} is *another* **Einstein B-coefficient**. (Note: It is also common to formulate this discussion using the energy density u_v instead of I_v.)

The number of transitions per second per cubic centimeter is found by multiplying each transition probability by the appropriate level population:

- $N_1 B_{12} I_v$ (for absorption)
- $N_2 A_{21}$ (for spontaneous emission)
- $N_2 B_{21} I_v$ (for stimulated emission)

Now for thermodynamic equilibrium, there will be a relation between the atoms and the radiation field. Thermodynamic equilibrium requires that there be no *net* change in the population of either level, meaning that the number of upward transitions must be equal to the number of downward transitions. As the atomic transition rates given above involve the intensity of the radiation, there is an definite connection between the radiation and the atoms, and we should be able to use this connection to deduce some properties of the photon field.

We begin by again stating the the number of upward transitions should equal the number of downward transitions, which provides our starting equation:

Eq. 4.7 $$N_1 B_{12} I_v = N_2 A_{21} + N_2 B_{21} I_v$$

This can be solved for the intensity:

Eq. 4.8 $$I_v = \frac{A_{21}/B_{21}}{\frac{N_1}{N_2}\frac{B_{12}}{B_{21}} - 1}$$

In thermodynamic equilibrium, the ratio of level populations is given by the Boltzmann equation:

Eq. 4.9 $$\frac{N_1}{N_2} = \frac{g_1}{g_2} e^{-(E_1 - E_2)/kT}$$

Note that as the temperature approaches infinity, the exponent diminishes to zero and the ratio of the level populations approaches the ratio of the statistical weights:

As $T \to \infty$, $\dfrac{N_1}{N_2} \to \dfrac{g_1}{g_2}$.

So we may interpret the ratio of statistical weights as the relative level population at infinite temperature.

We constructed our two-level atom such that its energy levels differ by the photon energy *hν*. Therefore, $E_2 - E_1 = h\nu$, and the Boltzmann equation is simplified:

Eq. 4.10 $$\frac{N_1}{N_2} = \frac{g_1}{g_2} e^{h\nu/kT}$$

Substituting this expression into Equation 4.8 gives the following:

Eq. 4.11 $$I_\nu = \frac{A_{21}/B_{21}}{\frac{g_1}{g_2}\frac{B_{12}}{B_{21}} e^{h\nu/kT} - 1}$$

For low frequencies ($h\nu/kT << 1$), the intensity of the radiation field is observed to follow the **Rayleigh-Jeans law**:

Eq. 4.12 $$I_\nu = \frac{2kT}{c^2}\nu^2$$

In order to get Equation 4.11 into this form, we must use the series expansion for the exponential:

Eq. 4.13 $$e^x \approx 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots \Rightarrow e^{\frac{h\nu}{kT}} \approx 1 + \frac{h\nu}{kT} + \frac{\left(h\nu/kT\right)^2}{2!} + \frac{\left(h\nu/kT\right)^3}{3!} + \ldots$$

Retaining only the first two terms of the expansion, substituting into Equation 4.11, and inserting the result into Equation 4.12 gives the following:

Eq. 4.14 $$I_\nu = \frac{A_{21}/B_{21}}{\frac{g_1}{g_2}\frac{B_{12}}{B_{21}}\left(1 + \frac{h\nu}{kT}\right) - 1}$$ and

Eq. 4.15 $$\frac{2kT}{c^2}\nu^2 = \frac{A_{21}/B_{21}}{\left(\frac{g_1}{g_2}\frac{B_{12}}{B_{21}} - 1\right) + \frac{g_1}{g_2}\frac{B_{12}}{B_{21}}\frac{h\nu}{kT}}$$

Now in order for the right side of Equation 4.15 to be a second-order polynomial in *ν*, we must have the quantity in parentheses equal to zero:

Eq. 4.16 $$g_1 B_{12} = g_2 B_{21}$$

With this requirement, we can find a relation between Einstein's A- and B-coefficients:

Eq. 4.17 $$\frac{A_{21}}{B_{21}} = \frac{2kT}{c^2}\nu^2\left(\frac{h\nu}{kT}\right) = \frac{2h\nu^3}{c^2}$$

Substituting the coefficient relations Equations 4.16 and 4.17 into Equation 4.11 yields the form of the intensity, good for all frequencies:

Eq. 4.18 $$I_\nu = \frac{2h\nu^3}{c^2}\frac{1}{e^{h\nu/kT}-1} = B_\nu(T)$$

Thus, our analysis of atomic transition probabilities produces the correct form of the Planck function! The three Einstein coefficients are related by two equations:

Eq. 4.19 $$B_{12} = \frac{g_2}{g_1}B_{21} \qquad \text{and}$$

Eq. 4.20 $$A_{21} = \frac{2h\nu^3}{c^2}B_{21}$$

From its original definition, we can see that the units of A_{21} should be [s^{-1}], and from its definition, we find the units of B_{21} (and B_{12}) should be [s^{-1}/(intensity units)], a combination that is not particularly useful to know.

What *is* important to know is that the Einstein coefficients – which relate to transition probabilities for a given set of atomic energy levels – are *atomic properties*, and thus, independent of temperature. This means that these relations should hold, even *without* thermodynamic equilibrium. Also, these coefficients can be applied to *any* transition between *any* two levels – not just the levels in our fictitious two-level atom.

The values of the coefficients depend on the atom and levels involved, and also on the type of transition. Order of magnitude estimates for the A-coefficient are given in Table 4.3.

Table 4.3: Transition probabilities and lifetimes

Transition	A_{21}	Lifetime $\approx {}^1/_{A_{21}}$
Electric Dipole	$\approx 10^8\ s^{-1}$	$\approx 10^{-8}$ s
Magnetic Dipole	$\approx 10^4\ s^{-1}$	$\approx 10^{-4}$ s
Electric Quadrupole	$\approx 10\ s^{-1}$	$\approx 10^{-1}$ s

Absorption/Emission Coefficients

In Chapter 2, we introduced the absorption coefficient κ_ν and the emission coefficient j_ν; how do these relate to the Einstein coefficients?

In Equation 2.25 we defined the spontaneous (mass) emission coefficient in terms of the energy emitted per gram of matter: $dE_\nu = j_\nu\, dt\, d\nu\, d\omega$.

We can now use dimensional analysis to write the amount of energy emitted per gram in terms of the Einstein A-coefficient:

Eq. 4.21 $$dE_\nu\left(\frac{\text{ergs}}{\text{g}}\right) = \left[\frac{h\nu(\text{ergs})}{4\pi(\text{st})}\frac{A_{21}}{(\text{s})}\frac{N_2}{(\text{cc})}\frac{1}{\rho}\left(\frac{\text{cc}}{\text{g}}\right)\frac{\phi(\nu)}{(\text{Hz})}\right] d\omega\,(\text{st})\,dt(\text{s})\,d\nu(\text{Hz})$$

Here the quantity $\phi(\nu)$ is the **line profile** – the strength of the emission as a function of frequency in the neighborhood of the transition frequency. (Although we have said that emission and absorption require a photon with a frequency *exactly* matching the energy difference between the two levels involved, it turns out that the energy levels are really not infinitely sharp; this permits a narrow range of photon frequencies to participate, which leads to the line profile. See Chapter 6.)

Comparing Equation 4.21 to Equation 2.25, it becomes apparent that we may identify the quantity in brackets in Equation 4.21 as j_ν:

Eq. 4.22 $$j_\nu = \frac{h\nu}{4\pi\rho} N_2 A_{21} \phi(\nu)$$

Similarly, we may relate the mass absorption coefficient to the appropriate Einstein B-coefficient:

Eq. 4.23 $$\kappa_\nu = \frac{h\nu}{4\pi\rho} N_1 B_{12} \phi(\nu)$$

What about stimulated emission? Where does it fit, now that we have accounted for both κ_ν and j_ν? Because stimulated emission is proportional to the intensity, it is easier to treat it as **negative absorption**. Then we can modify Equation 4.23 as follows:

Eq. 4.24 $$\kappa_\nu = \frac{h\nu}{4\pi\rho} \phi(\nu)\left(N_1 B_{12} - N_2 B_{21}\right)$$

This is the absorption coefficient, corrected for stimulated emission. What is the magnitude of the correction?

We can transform this expression by substituting a Boltzmann equation for N_2 and using Equation 4.19 to replace B_{21}:

Eq. 4.25 $$N_2 B_{21} = \left(N_1 \frac{g_2}{g_1} e^{-h\nu/kT} \right)\left(\frac{g_1}{g_2} B_{12} \right) = N_1 B_{12} e^{-h\nu/kT}$$

The absorption coefficient is then as follows:

Eq. 4.26 $$\kappa_\nu = \frac{h\nu}{4\pi\rho} \phi(\nu) N_1 B_{12}\left(1 - e^{-h\nu/kT}\right)$$

The exponential term can be estimated for typical values ($\lambda = 4000$ and $T = 10{,}000$):

Eq. 4.27 $$\frac{h\nu}{kT} = \frac{hc}{\lambda kT} \approx \frac{12400}{(4000)(8.6e-5)(10000)} \approx 3.6 \Rightarrow e^{-x} \approx 0.03$$

So stimulated emission typically reduces absorption by a few percent in the visible.

It can be seen from Equations 4.22 and 4.26 that the absorption coefficient depends on the population of the lower level while the emission coefficient depends on the population of the upper level, both of which can be determined from the Boltzmann equation and are thus

dependent on temperature. But there is another important process that gives the electrons additional options on ways to spend their time and energy.

Ionization

In addition to hopping around from level to level within the atom, an electron may also pack its bags and leave altogether, in a process known as ionization. This could be thought of as propelling the electron into a super-excited state, one whose population – which we would like to know – could perhaps be described by a Boltzmann equation of sorts. First, let us decide on some terminology.

For simplicity, let us begin with a neutral atom in the ground state (X_0); we will then let it ionize, which creates two particles: a free electron (e^-) and an ion (X_0^+ – also in the ground state). These three particles will have number densities, energies, and statistical weights as shown in Table 4.4.

Table 4.4: Ionization terminology

Species:	Ground State Neutral Atom	⇒	Ground State Ion	+	Free Electron
Reaction:	X_0	⇒	X_0^+	+	e^-
Number Density:	N_0	⇒	N_0^+	+	N_e
Energy:	0	⇒	χ	+	$^1/_2 m_e \mathrm{v}^2$
Statistical Weight:	g_0	⇒	g_0^+	+	g_e

Here the term χ represents the **ionization energy** for this atom – the difference in energy between the ground state ion and the ground state neutral. The free electron may have kinetic energy of its own, and this may be any amount, determined by the electron's velocity. The total energy difference between the ion-electron combination and the neutral atom is then as follows:

Eq. 4.28 $$\Delta E = \chi + {}^1/_2 m_e \mathrm{v}^2$$

The overall statistical weight of the ionized state (g) is the product of the statistical weights of the ion and the electron

Eq. 4.29 $$g = g_o{}^+ g_e$$

We will now let $\Delta N_0^+(\mathrm{v})$ be the number density of ground state ions with electrons having speeds in the range $\mathrm{v} \rightarrow \mathrm{v} + \Delta\mathrm{v}$. The statistical weight of these ion-electron combinations is then as follows:

Eq. 4.30 $$\Delta g = g_o{}^+ \Delta g_e + g_e \Delta g_o{}^+$$

(The second term disappears because the statistical weight of the ground state ion is fixed; we are only varying the electron speed.)

If the electron density is given by N_e, then the volume containing one electron – expressed here as a cube of sides Δx, Δy, and Δz – will be the inverse of this quantity:

Eq. 4.31 $$\Delta V = \Delta x\, \Delta y\, \Delta z = {}^1/_{N_e}$$

Now the **Heisenberg uncertainty principle** (Equation 4.32) tells us that we cannot know both the position and the momentum of a particle *precisely*; there will be an uncertainty associated with each of these quantities, and these uncertainties are related:

Eq. 4.32 $\Delta x\, \Delta p_x \geq h$ (and similarly for *y* and *z*)

(Note: Depending on the ultimate goal, the right side of Equation 4.32 may be *h*, *ħ*, $h/2$, *or* $\hbar/2$, as needed.)

We may now combine the three uncertainly principle equations for *x*, *y*, and *z* to obtain a statement about the minimum volume an electron may occupy (ΔV_{min}):

Eq. 4.33 $$\Delta V = \Delta x\, \Delta y\, \Delta z \quad \geq \quad \frac{h^3}{\Delta p_x \Delta p_y \Delta p_z} = \Delta V_{\min}$$

That is, the electrons in a given gas of electron density N_e each have available an average volume ΔV; the smallest volume each electron could possibly squeeze into (given the current state of the gas) is ΔV_{min}. This means that the number of different ways that an electron may exist in the gas (the statistical weight) is the current volume divided by the minimum volume:

Eq. 4.34 $$\Delta g_e = 2\frac{\Delta V}{\Delta V_{\min}}$$

The factor of 2 in this equation comes, of course, from the spin. Each minimum volume may contain two electrons, but only if they have opposite spin.

We can substitute into Equation 4.34 to find this statistical weight in terms of the electron momentum:

Eq. 4.35 $$\Delta g_e = 2\frac{\Delta V}{\Delta V_{\min}} = 2\frac{1/N_e}{h^3/\Delta p_x \Delta p_y \Delta p_z} = \frac{2\Delta p_x \Delta p_y \Delta p_z}{h^3 N_e}$$

If we now assume an isotropic distribution of speeds, and hence, momenta, we may write the following:

Eq. 4.36 $$\Delta p_x\, \Delta p_y\, \Delta p_z = 4\pi\, p^2 \Delta p = 4\pi\, m_e{}^3\, \mathrm{v}^2 \Delta \mathrm{v}$$

This allows us to write the statistical weight of the free electrons (associated with the ions described above) as follows:

Eq. 4.37 $$\Delta g_e = \frac{8\pi\, m_e^3 \mathrm{v}^2 \Delta \mathrm{v}}{h^3 N_e}$$

At last we are ready to write a Boltzmann equation to relate the population of ions with electrons in a particular speed range to the population of ground state neutral atoms:

Eq. 4.38 $$\frac{\Delta N_o^+(\mathrm{v})}{N_o} = \frac{\Delta g}{g_o} e^{-\Delta E/kT} = \frac{g_o^+}{g_o}\frac{8\pi\, m_e^3}{h^3 N_e} e^{-\left(\chi + \frac{1}{2} m_e \mathrm{v}^2\right)/kT} \mathrm{v}^2 \Delta \mathrm{v}$$

Of course, we are really not concerned with a particular speed range for the electrons; we just want to know the relative populations of ions and neutrals. Thus, we will integrate over all speeds, employing the substitution of the dummy variable $x \equiv \sqrt{\frac{m_e}{2kT}}\mathrm{v}$:

Eq. 4.39 $$\frac{N_o^+}{N_o} = \frac{8\pi m_e^3}{h^3 N_e}\frac{g_o^+}{g_o} e^{-\chi/kT}\left(\frac{2kT}{m_e}\right)^{3/2}\int_0^\infty x^2 e^{-x^2} dx$$

The integral has a value of $\sqrt{\pi}/4$, and this gives the following result:

Eq. 4.40 $$\frac{N_o^+ N_e}{N_o} = \frac{2g_o^+}{g_o}\left(\frac{2\pi m_e kT}{h^2}\right)^{3/2} e^{-\chi/kT}$$

This equation is one form of the **Saha equation**.

Saha Equation

In general, the Saha equation relates the population of ions to that of neutrals, but there are several ways to do this, and several forms of the Saha equation as well. The one given above relates ground state ions to ground state neutrals.

We can broaden this expression to include *all* ions and *all* neutrals by noting the following relations. The ratio of ground state neutrals to *all* neutrals is given by $\frac{N_o}{N} = \frac{g_o}{U(T)}$, and similarly, the ratio of ground state ions to *all* ions is given by $\frac{N_o^+}{N^+} = \frac{g_o^+}{U^+(T)}$. We can substitute these expressions into Equation 4.40 to obtain a Saha equation that relates the total number density of ions (N^+) to the total number density of neutrals (N):

Eq. 4.41 $$\frac{N^+ N_e}{N} = \frac{2U^+(T)}{U(T)}\left(\frac{2\pi m_e kT}{h^2}\right)^{3/2} e^{-\chi/kT}$$

But the Saha equation is even more general, as we may use it to relate the number densities of any two successive ionization states (i and $i+1$):

Eq. 4.42 $$\frac{N_{i+1} N_e}{N_i} = \frac{2U_{i+1}(T)}{U_i(T)}\left(\frac{2\pi m_e kT}{h^2}\right)^{3/2} e^{-\chi_i/kT}$$

There are other ways to write the Saha equation that employ **partial pressures** rather than number densities of the different components. Recall the **ideal gas law** $PV = n\Re T$ (where n is *not* a quantum number but the number of moles of the gas, and $\Re$ is the gas constant*); this can be rewritten, using **Avogadro's number** (N_A)* as follows:

* See Appendix: Constants

Eq. 4.43 $$P = \frac{n}{V}\Re T = \frac{nN_A}{V}\frac{\Re}{N_A}T = NkT$$

Thus, the **electron pressure** is $P_e = N_e kT$; similar expressions can be written for the other partial pressures. In terms of partial pressures, the Saha equation is then as follows (note the extra power of kT):

Eq. 4.44 $$\frac{P_{i+1}P_e}{P_i} = \frac{N_{i+1}P_e}{N_i} = \frac{2U_{i+1}(T)}{U_i(T)}\left(\frac{2\pi m_e}{h^2}\right)^{3/2}(kT)^{5/2}e^{-\chi_i/kT}$$

Equations 4.42 and 4.44 are general equations, valid for any species in local thermodynamic equilibrium. They require input of properties – ionization potential, excitation energies, and statistical weights – that are specific to the atom involved. These are found in tables in the literature:

	Cox (2000)	Allen (1973)	Novotny (1973)
Ionization potentials:	table 3.5	section 16	table 3-4
Statistical weights:	table 3.3	section 15	table 3-2A
Partition functions:	table 3.3	section 15	table 3-2B

The Hydrogen Atom

In real life, atoms do not focus separately on either excitation or ionization; rather they tackle both processes at the same time. In a given gas, some of the neutral atoms will be in the ground state while others will be excited, and some of the atoms will have been ionized, with some of these ions being excited as well. Although this may sound complicated, we now have the tools to predict how the atoms will configure themselves for a given set of gas properties – as long as we can assume that the gas has reached thermodynamic equilibrium (which is usually a good assumption, at least as a first approximation).

As an example of how this works, let us consider a gas of hydrogen atoms and inquire how these atoms will be distributed among the various possible energy states for a given set of conditions. We will assume that both excitation and ionization processes have had sufficient time to reach a state of equilibrium.

We must first establish our terminology. We will allow the neutral hydrogen atoms to be distributed over the different energy levels ranging from $n = 1$ (the ground state) to $n = m$ (a practical upper limit). As ionized hydrogen has no bound electron, we need only consider one ionized state. Let the associated number densities be denoted by N_1 , N_2 , N_3 , ... N_m , and N^+. As we will be calculating a number of ratios, it will be convenient to define these as well:

Let $\beta_2 \equiv N_2/N_1$, $\beta_3 \equiv N_3/N_1$, $\beta_m \equiv N_m/N_1$, and $\beta^+ \equiv N^+/N_1$.

The total number density of hydrogen atoms/ions is found by adding the number densities of all of the neutrals and ions:

Eq. 4.45 $$N_T = N_1 + N_2 + N_3 + ... + N_m + N^+ = N_1(1 + \beta_2 + \beta_3 + ... + \beta_m + \beta^+)$$

The β_n ratios can be found from Boltzmann equations similar to Equation 4.3, using the substitutions $g_n = 2n^2$ and $\Delta E_n = E_n - E_1 = E_0\,(1 - {}^1/_{n^2})$:

Eq. 4.46 $$\beta_n = \frac{N_n}{N_1} = \frac{g_n}{g_1} e^{-\Delta E_n/kT} = n^2 e^{-\frac{E_o}{kT}\left(1-\frac{1}{n^2}\right)}$$

The β^+ ratio can be determined by a Saha equation, based on either electron pressure or electron number density:

Eq. 4.47 $$\beta^+ = \frac{N^+}{N_1} = \frac{2g_+}{g_1}\left(\frac{2\pi m_e k}{h^2}\right)^{3/2} k\frac{T^{5/2}}{P_e} e^{-\chi/kT}$$ or

Eq. 4.48 $$\beta^+ = \frac{N^+}{N_1} = \frac{2g_+}{g_1}\left(\frac{2\pi m_e k}{h^2}\right)^{3/2} \frac{T^{3/2}}{N_e} e^{-\chi/kT}$$

We now make a few simplifications, noting that the constants E_0 and χ are the same for hydrogen; thus we may define $E_0/k = \chi/k \equiv B$ (Pierce's B-coefficient), which has a value of about 157,800. In both Saha equations we find another constant $\left(\left(\frac{2\pi m_e k}{h^2}\right)^{3/2} \equiv A\right)$ (Pierce's A-coefficient), which has a value of about 2.4147e15; this can be combined with k in Equation 4.47 to form $Ak \approx 0.33338$ (another of Pierce's A-coefficients).

We also need statistical weights for the Saha equations. We already know that the ground state hydrogen atom has statistical weight $g_1 = 2(1)^2 = 2$, but what is the statistical weight of a hydrogen ion (g^+)? This turns out to be relatively easy, as there is only *one* way to have a proton, making $g^+ = 1$; then $2g^+/g_1 = 2(1)/2 = 1$. With these simplifications, the Boltzmann and Saha equations for hydrogen are as follows:

Eq. 4.49 $$\beta_n = n^2 e^{-\frac{B}{T}\left(1-\frac{1}{n^2}\right)}$$

Eq. 4.50 $$\beta^+ = \frac{Ak}{P_e} T^{5/2} e^{-B/T}$$

Eq. 4.51 $$\beta^+ = \frac{A}{N_e} T^{3/2} e^{-B/T}$$

It is now a simple matter to use these equations to find the population of any level relative to the ground state of neutral hydrogen. While the temperature is sufficient to determine the populations of excited states, the population of ions depends on both temperature and pressure (or density), as can be seen from the presence of the electron pressure in Equation 4.50 and the electron density in Equation 4.51. One of these – whichever is more convenient – will be used to calculate ionization.

As an example of how excitation and ionization vary with temperature, consider the data in Table 4.5, for which an electron pressure of 10 (dynes/cm^2) has been used, and excitation states up to m = 14 have been included.

Table 4.5: Level populations of hydrogen (relative to ground state) for P_e = 10

n	T=6000	T=8000	T=10000	T=12000
1	1	1	1	1
2	1.09e–8	1.50e–6	2.90e–5	2.08e–4
3	6.33e–10	2.19e–7	7.29e–6	7.55e–5
4	3.13e–10	1.49e–7	6.02e–6	7.08e–5
5	2.71e–10	1.49e–7	6.59e–6	8.23e–5
6	2.83e–10	1.69e–7	7.83e–6	1.01e–4
7	3.17e–10	1.99e–7	9.48e–6	1.25e–4
8	3.65e–10	2.36e–7	1.15e–5	1.53e–4
9	4.24e–10	2.80e–7	1.38e–5	1.85e–4
10	4.92e–10	3.31e–7	1.64e–5	2.22e–4
11	5.69e–10	3.86e–7	1.93e–5	2.62e–4
12	6.54e–10	4.48e–7	2.25e–5	3.07e–4
13	7.47e–10	5.15e–7	2.60e–5	3.55e–4
14	8.48e–10	5.88e–7	2.98e–5	4.08e–4
+	0.00035	0.51785	46.7490	1023.105
Total	1.00035	1.51786	47.7492	1024.108
% Ionized	0.04%	34.12%	97.91%	99.90%

Several points should be noted:

- Excitation (of every level) increases with temperature. As more energy is made available to the atoms (by way of increased kinetic energy of gas particles), they tend to store more of this energy as excitation energy.
- Similarly, ionization increases with temperature as well, as kinetic energy is stored as ionization energy.
- All of the excited state populations for hydrogen are *small* relative to the ground state; most neutral hydrogen atoms will be in the ground state, due to the large energy difference between the first two levels.
- The degree of ionization increases dramatically with temperature, such that the ion population changes from *insignificant* to *dominant* in only a few thousand degrees. Consequently, the excitation of the few remaining neutral atoms becomes less important.

Figure 4.2 shows the ionization of hydrogen as a function of temperature; excited states have been ignored in this calculation of the percentages.

Figure 4.2: Hydrogen ionization at $P_e = 10$

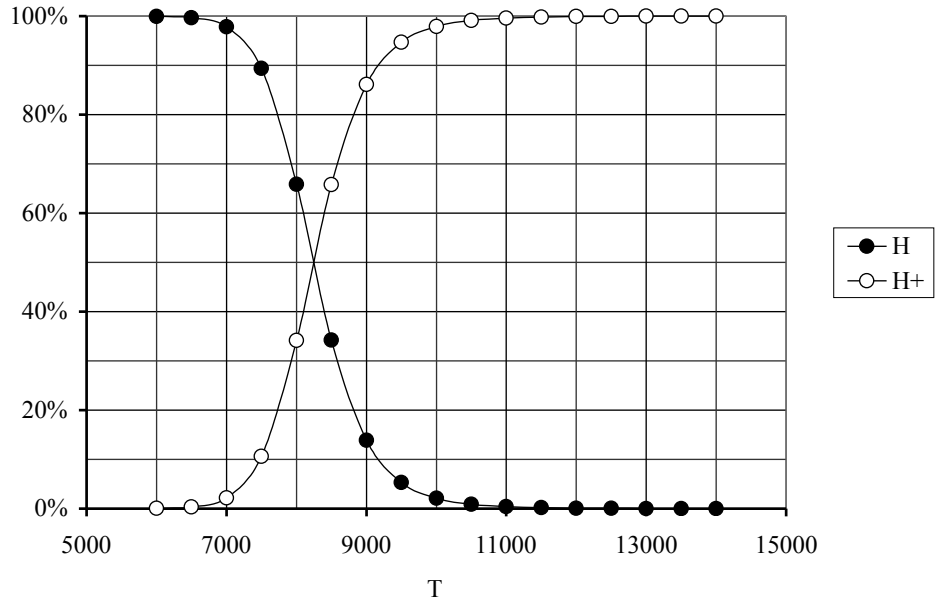


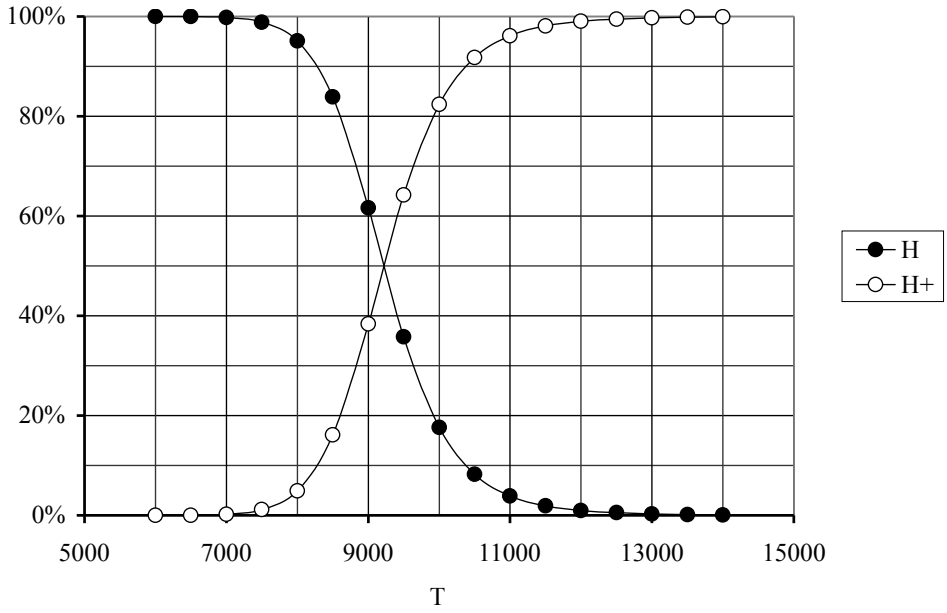
Table 4.6: Level populations of hydrogen (relative to ground state) for T = 9000

n	P_e=1	P_e=10	P_e=100	P_e=1000
1	1	1	1	1
2	7.78e–6	7.78e–6	7.78e–6	7.78e–6
3	1.53e–6	1.53e–6	1.53e–6	1.53e–6
4	1.16e–6	1.16e–6	1.16e–6	1.16e–6
5	1.22e–6	1.22e–6	1.22e–6	1.22e–6
6	1.42e–6	1.42e–6	1.42e–6	1.42e–6
7	1.70e–6	1.70e–6	1.70e–6	1.70e–6
8	2.04e–6	2.04e–6	2.04e–6	2.04e–6
9	2.44e–6	2.44e–6	2.44e–6	2.44e–6
10	2.89e–6	2.89e–6	2.89e–6	2.89e–6
11	3.40e–6	3.40e–6	3.40e–6	3.40e–6
12	3.95e–6	3.95e–6	3.95e–6	3.95e–6
13	4.55e–6	4.55e–6	4.55e–6	4.55e–6
14	5.21e–6	5.21e–6	5.21e–6	5.21e–6
+	62.218	6.2218	0.62218	0.06222
Total	63.218	7.2218	1.62222	1.06226
% Ionized	98.42%	86.15%	38.35%	5.86%

The above data was produced at a constant electron pressure. If we hold temperature constant and vary the electron pressure, we obtain the *pressure* dependence of ionization (excitation being independent of pressure). We can predict this relation by examining the ionization reaction ($X \rightleftharpoons X^+ + e^-$); because the forward reaction transforms one particle into two,

increasing the pressure will push the reaction backwards, towards the neutral atom. This result can be seen in Table 4.6.

Figure 4.3: Hydrogen ionization at $P_e = 100$



As electron pressure increases, the degree of ionization decreases. And as we shall soon see, this behavior will be true for the total gas pressure as well. Figure 4.3 (compare with Figure 4.2) shows the effect on ionization of increasing the electron pressure by a factor of 10, to 100. As can be seen, the temperature at which the gas is 50% ionized has increased by about 1000 K.

The principal variables for ionization are temperature and pressure (or density), but the atomic parameters (statistical weight and ionization potential) play a role as well. Of these, the more important is ionization potential, due to its position in the exponential. Elements with low ionization potentials are considerably easier to ionize, and this behavior is observed in stellar spectra.

Gas-in-a-box Problem

In the above ionization calculations, we specified the electron pressure (or density) and determined the resulting degree of ionization. But this is not how the stars work. Stars do not have a dial labeled 'Electron Pressure'; instead, they must produce their own electrons by the ionization process, which of course depends on what the electron pressure is. Stars seem to be very adept at reaching ionization equilibrium, and we should be able to figure it out as well. Let us begin with a simple example: the gas-in-a-box problem.

Suppose that we have a box in which we put atoms of a particular elemental gas, such as atomic hydrogen. We then heat the gas to some equilibrium temperature and inquire as to the distribution of particles within. (We may assume (or require) that the hydrogen *not* form molecules or any other species other than the neutral atom and the ion.) This time we do not know the electron pressure (or density), but we do know the volume of the box and how many

atoms we have put into it, giving us the total number density of hydrogen nuclei (N_H). What is the degree of ionization?

By assuming that the hydrogen nuclei must exist as either neutrals ($N°$) or ions (N^+), we are requiring that $N_H = N° + N^+$. We are also assuming that the only source of electrons is the hydrogen itself, which means that the electron density must equal the number density of ions: $N_e = N^+$. Combining these two gives an expression for the number density of neutrals in terms of a given quantity and an unknown:

Eq. 4.52 $$N° = N_H - N_e$$

We will simplify our analysis by assuming that partition functions can be approximated by the ground state statistical weights: $U(T) \approx g_0$, $U^+(T) \approx g^+$, etc. The appropriate Saha equation is then as follows (from Equation 4.40):

Eq. 4.53 $$\frac{N^+}{N^o} = \frac{N_e}{N_H - N_e} = \frac{2g^+}{g_o N_e}\left(\frac{2\pi m_e kT}{h^2}\right)^{3/2} e^{-\chi/kT}$$

Recalling that $2g^+/g_0 = 1$ for hydrogen, the variable N_e can then be collected on the left side, leaving a function of temperature on the right:

Eq. 4.54 $$\frac{N_e^2}{N_H - N_e} = AT^{3/2}e^{-B/T} = f(T)$$

This is clearly a quadratic equation:

Eq. 4.55 $$N_e^2 + f(T)N_e - f(T)N_H = 0 \quad \Rightarrow \quad ax^2 + bx + c = 0$$

And its solution is given by the well-known quadratic formula:

Eq. 4.56 $$x = -\left(\frac{b}{2a}\right) \pm \sqrt{\left(\frac{b}{2a}\right)^2 - \frac{c}{a}} \quad \Rightarrow \quad N_e = -\frac{f(T)}{2} + \sqrt{\left(\frac{f(T)}{2}\right)^2 + f(T)N_H}$$

Here we have rejected the negative root (as the lead term is already negative, and we demand a positive value for N_e). It should be readily apparent that no matter what the temperature, the electron density should increase as the number density of hydrogen increases. But what about the ionization degree? The fraction of the gas that is ionized is simply the ratio of ions to total hydrogen, which is equal to N_e/N_H.

Eq. 4.57 $$\frac{N_e}{N_H} = -\frac{f(T)}{2N_H} + \sqrt{\left(\frac{f(T)}{2N_H}\right)^2 + \frac{f(T)}{N_H}}$$

This ratio decreases with increasing N_H, for any temperature. Table 4.7 shows how the ionization degree varies with N_H for a temperature of 7000:

Table 4.7: Hydrogen gas-in-a-box for T = 7000

log N_H	N_e	% Ionized
9	9.96e+8	99.6%
10	9.60e+9	96.0%
11	7.53e+10	75.3%
12	3.78e+11	37.8%
13	1.40e+12	14.0%
14	4.67e+12	4.7%
15	1.50e+13	1.5%
16	4.78e+13	0.5%

As predicted, high density suppresses ionization. We should expect that denser stellar atmospheres should be less ionized (assuming equal temperatures), and this is indeed the case.

Gas Mixtures

Although most stars are composed primarily of hydrogen, they do contain other elements as well. Some of these elements have lower ionization potentials than hydrogen, and thus they will ionize more readily. We now investigate how the ionization of one element will affect the ionization of others.

Consider a mixture of two gases, elements *X* and *Y*, in a box, at densities N_x and N_y. For simplicity, assume that partition functions can be approximated by statistical weights and let $G \equiv 2g^+/g_0$ for each element. We can then write the ionization reaction and resulting Saha equation for each element:

Eq. 4.58 $$X \rightleftharpoons X^+ + e^- \quad \Rightarrow \quad \frac{N_x^+}{N_x^o} = G_x \frac{A}{N_e} T^{3/2} e^{-\chi_x/kT}$$

Eq. 4.59 $$Y \rightleftharpoons Y^+ + e^- \quad \Rightarrow \quad \frac{N_y^+}{N_y^o} = G_y \frac{A}{N_e} T^{3/2} e^{-\chi_y/kT}$$

The key point to note concerns the electron density N_e, which appears in both equations. This quantity includes electrons from *both* elements; that is, $N_e = N_{ex}+N_{ey} = N_x^+ + N_y^+$. However, evaluation of *either* Saha equation requires knowledge of the value of N_e, which in turn depends on the results of *both* Saha equations. This dilemma has a relatively simple solution:

1. *Assume* a value for N_e.
2. Calculate the ratios $\frac{N_x^+}{N_x^o}$ and $\frac{N_y^+}{N_y^o}$ using the Saha equations.
3. Calculate N_x^+ from $\frac{N_x^+}{N_x^o}$ and N_x, and N_y^+ from $\frac{N_y^+}{N_y^o}$ and N_y.
4. Calculate a new value of $N_e = N_x^+ + N_y^+$.

5. Go to Step 2 and iterate until the process converges on one value of N_e.

This process is illustrated by an example involving a gas of the elements sodium and magnesium. The necessary equations are relatively simple:

Eq. 4.60 $$\beta_i = G_i \frac{A}{N_e} T^{3/2} e^{-\chi_i/kT} = \frac{f_i(T)}{N_e}$$

Eq. 4.61 $$N_i^+ = \left(\frac{\beta_i}{1+\beta_i} \right) N_i$$

Eq. 4.62 $$N_e = N_x^+ + N_y^+$$

The constants needed are A = 2.4147e+15 and k = 8.6173e–5, and the atomic parameters as found in the literature are given in Table 4.8.

Table 4.8: Atomic properties for the mixture-in-a-box problem

Element	Na (X)	Mg (Y)
$G = 2g^+/g_0$	1	4
χ (eV)	5.1391	7.6462

We can try to predict which element will have a higher degree of ionization by looking at the atomic properties. The statistical weights give an edge to magnesium, but the ionization potentials favor sodium; which of these will prevail depends on the temperature.

We now choose T = 6000, giving $f(T)$ values of 5.413e+16 for sodium and 1.697e+15 for magnesium. Because $f(T)$ is proportional to the ion/neutral ratio, these values tell us that a greater fraction of sodium should be ionized.

The last things needed to solve the problem are the number densities of each element. For simplicity, we will assume equal number densities: $N_x = N_y$ = 1.2e+16. We next make an initial guess of the electron density: N_e = 5e+15. Results of the ensuing calculations are shown in Table 4.9.

Table 4.9: Iterative solution of the mixture-in-a-box problem

N_e	β_x	β_y	N_x^+	N_y^+	$N_e = N_x^+ + N_y^+$
5e+15	10.83	0.3394	1.099e+16	3.040e+15	1.4026e+16
1.4026e+16	3.859	0.1210	9.531e+15	1.295e+15	1.0826e+16
1.0826e+16	5.000	0.1567	1.000e+16	1.626e+15	1.1626e+16
1.1626e+16	4.656	0.1459	9.878e+15	1.528e+15	1.1407e+16
1.1407e+16	4.746	0.1488	9.911e+15	1.554e+15	1.1465e+16
1.1465e+16	4.721	0.1480	9.903e+15	1.547e+15	1.1449e+16
1.1449e+16	4.728	0.1482	9.905e+15	1.549e+15	1.1454e+16
1.1454e+16	4.726	0.1481	9.904e+15	1.548e+15	1.1453e+16
1.1453e+16	4.726	0.1482	9.904e+15	1.548e+15	1.1453e+16

The final row gives the solution, for which the electron density is indeed the sum of the two ion densities. Comparison of these ion densities with the elemental number densities shows that the sodium is 82.54% ionized and the magnesium is only 12.90% ionized – which is consistent with our prediction.

This problem can sometimes be simplified if the ionization potentials are quite different from each other. For example, if χ_x = 12 eV and χ_y = 5 eV, then *Y* will normally ionize much more easily than *X*, such that we may presume that *Y* will contribute the majority of the electrons in the gas, assuming comparable abundances and moderate temperatures. We may then approximate the electron density by solving the *Y*-in-a-box problem:

Eq. 4.63 $$N_e \approx N_{ey} = -\frac{f_y(T)}{2} + \sqrt{\left(\frac{f_y(T)}{2}\right)^2 + f_y(T) N_y}$$

We next insert this value into the Saha equation for *X* to solve for β_x and hence, N_{ex}. One should then reaffirm that N_{ex} is indeed much less than N_{ey}. If not, use the sum $N_e = N_{ex} + N_{ey}$ as an initial guess and proceed with the first method.

Of course, real stellar atmospheres have many different elements, but the basic approach will still work. However, most elements have multiple ionization stages, shedding more than one electron at higher temperatures. How can we deal with this problem?

Multiple Ionization Stages

Let us again consider the single gas-in-a-box problem, only this time the gas will be allowed to ionize twice. There will be three atomic/ionic species (X, X^+, X^{++}) along with electrons. We will need four number densities ($N°$, N^+, N^{++}, N_e), three partition functions ($U°$, U^+, U^{++}), and the first two ionization potentials for the element (χ_1, χ_2). Equipped with these, we can now write two Saha equations:

Eq. 4.64 $$\frac{N^+}{N^o} = \frac{2U^+(T)}{U^o(T)} \left(\frac{2\pi m_e kT}{h^2}\right)^{3/2} \frac{e^{-\chi_1/kT}}{N_e} = \frac{f_1(T)}{N_e}$$

Eq. 4.65 $$\frac{N^{++}}{N^+} = \frac{2U^{++}(T)}{U^+(T)} \left(\frac{2\pi m_e kT}{h^2}\right)^{3/2} \frac{e^{-\chi_2/kT}}{N_e} = \frac{f_2(T)}{N_e}$$

We may choose *T* and calculate the two temperature-dependent functions:

Eq. 4.66 $$f_1(T) = \frac{2U^+(T)}{U^o(T)} A T^{3/2}\, e^{-\chi_1/kT}$$

Eq. 4.67 $$f_2(T) = \frac{2U^{++}(T)}{U^+(T)} A T^{3/2}\, e^{-\chi_2/kT}$$

This allows us to write the number densities of the ions in terms of the number densities of the neutrals and the electrons:

Eq. 4.68 $$N^+ = f_1(T)\, N°/N_e$$

Eq. 4.69 $N^{++} = f_2(T)\, N^{+}/N_e = f_1(T)\, f_2(T)\, N^{\circ}/N_e^{\,2}$

The total number density of *all* nuclei of X is the sum of the neutrals and ions:

Eq. 4.70 $N_x = N^{\circ} + N^{+} + N^{++} = N^{\circ} + f_1(T)\, N^{\circ}/N_e + f_1(T)\, f_2(T)\, N^{\circ}/N_e^{\,2}$

Eq. 4.71 $N_x = N^{\circ}\{1 + [f_1(T)/N_e][1 + f_2(T)/N_e]\}$

Now given a value of N_x, we can guess N_e and calculate N°, and from this find N^{+} and N^{++}. The new value of N_e is now given by $N_e = N^{+} + 2N^{++}$, because each of the doubly ionized atoms contributes *two* electrons to the gas. Sufficient iteration then yields a solution.

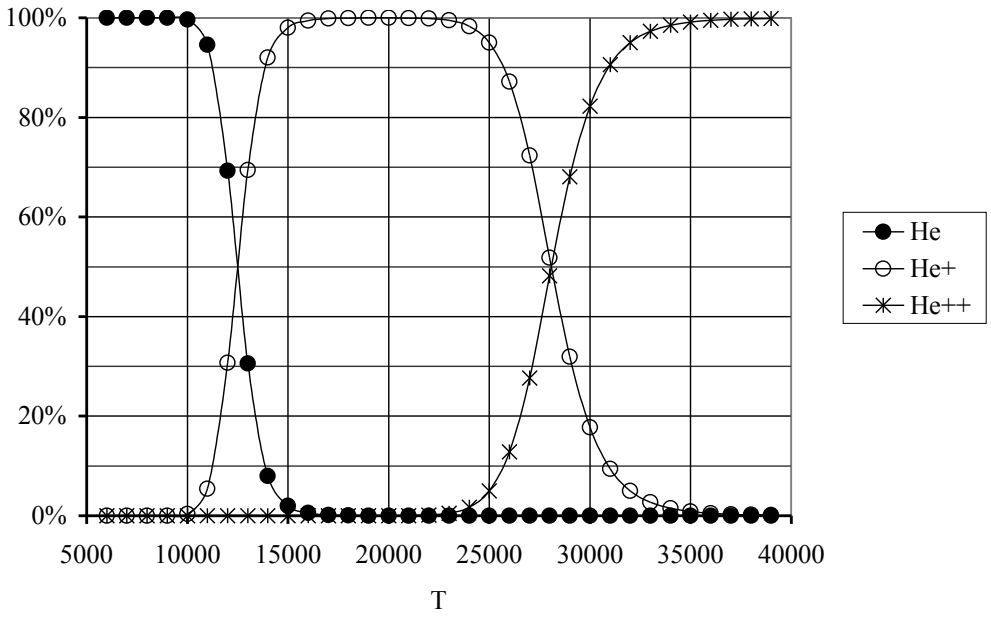
Results of such a calculation for helium are presented in Table 4.10 and Figure 4.4. Ionization stages have been determined for a helium number density of 1e+12 over temperatures ranging from 9000 to 36,000 K.

Table 4.10: Ionization equilibrium for helium at $N_{He} = 1 \times 10^{12}$

T	$f_1(T)$	$f_2(T)$	N_e	N°	N^{+}	N^{++}	$\Rightarrow N_e$
9000	1.4e+8	6.9e–10	1.2e+10	9.9e+11	1.2e+10	6.9e–10	1.2e+10
12000	6.0e+11	4.4e–2	5.3e+11	4.7e+11	5.3e+11	4.4e–2	5.3e+11
15000	9.7e+13	2.3e+3	9.9e+11	1.0e+10	9.9e+11	2.3e+3	9.9e+11
18000	3.0e+15	3.4e+6	1.0e+12	3.3e+8	1.0e+12	3.4e+6	1.0e+12
21000	3.7e+16	6.4e+8	1.0e+12	2.7e+7	1.0e+12	6.4e+8	1.0e+12
24000	2.5e+17	3.4e+10	1.0e+12	4.1e+6	9.7e+11	3.1e+10	1.0e+12
27000	1.1e+18	7.5e+11	1.4e+12	7.9e+5	6.4e+11	3.6e+11	1.4e+12
30000	3.7e+18	9.1e+12	1.8e+12	8.4e+4	1.7e+11	8.3e+11	1.8e+12
33000	1.0e+19	7.1e+13	2.0e+12	5.1e+3	2.7e+10	9.7e+11	2.0e+12
36000	2.4e+19	4.0e+14	2.0e+12	4.0e+2	4.9e+9	1.0e+12	2.0e+12

Generally only one or two ionization stages of a particular element are present in significant amounts at a given temperature – essentially no neutral helium remains at the temperatures where the second ionization begins. Thus, a model stellar atmosphere is unlikely to require *all* possible ionization states of every element; only those that play a significant role at the temperature of interest need be included – unless of course it is the trace amounts that are of interest.

Figure 4.4: Ionization of helium at $N_{He} = 1 \times 10^{12}$



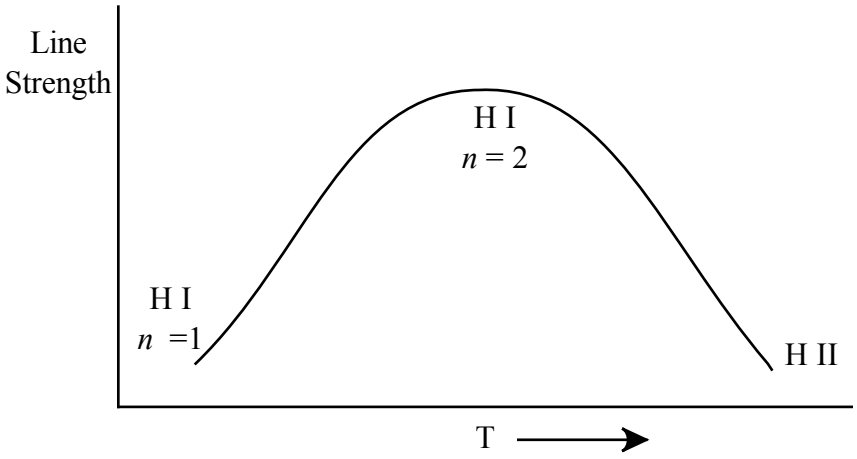
Effect on Transitions

The populations of different atomic/ionic species depend on the temperature and density (or pressure) in the atmosphere, via the Saha equation.

Energy level populations within a species depend on the temperature, via the Boltzmann equation.

A given absorption transition depends on the supply of absorbers (atoms/ions with electrons in the lower level of the transition), which depends on the temperature and density (or pressure) of the atmosphere.

Figure 4.5: Balmer line strength as a function of temperature



Each transition will have certain temperature-density combinations that are more favorable than others, with temperature being the more important variable. In general, the absorption **line**

strength will rise and fall over a range of stellar temperatures as lower excitation levels first become populated by higher temperatures and then depleted by ionization, as shown in Figure 4.5.

Although the peak line strength is labeled H I, $n = 2$, this does not mean that *every* neutral hydrogen is in the second level. As we have seen, only a tiny fraction of the neutrals will be so positioned, ready to do Balmer absorption; however, there are so many hydrogen atoms in most stars that even small fractions of them can produce noticeable results.

The temperature associated with the maximum line strength for a species depends on the atomic properties for that species and also on the density(or pressure) of the gas. In the next chapter we will step back and look at the big picture of how the many different atomic and ionic species combine to produce the variety of stellar spectra we observe.

CHAPTER 5: Stellar Spectra

The study of spectra began in 1666 when Isaac Newton passed sunlight through a prism, revealing the rainbow of colors that make up white light. In 1802, William Wollaston discovered that the spectrum of the Sun contained many dark lines; these lines were studied further and labeled by Joseph von Fraunhofer around 1814. In 1859, Gustav Kirchhoff performed experiments that finally explained how different types of spectra – including those with dark lines – arise from different types of light sources; his results are given in his three laws of spectral analysis.

Kirchhoff's Laws of Spectral Analysis

Kirchhoff's first law involves an opaque source, which could be a solid, a liquid, or a gas at high pressure. When such a source is heated to incandescence and the radiation is examined through a spectroscope, the spectrum shows light at all wavelengths – a **continuous spectrum**. The light is not uniformly intense at each wavelength; rather, the intensity varies slowly with wavelength, with the pattern depending on the temperature of the source. (We may generally regard this as a blackbody spectrum.)

For Kirchhoff's second law, the source is a hot, optically thin gas (a gas at low pressure). The spectrum of radiation from this source presents light only at certain wavelengths, a series of bright lines on an essentially dark background. This is called a **bright line spectrum** or an **emission spectrum**.

In Kirchhoff's third law, light from a hot opaque source is passed through a cooler, thin gas before being viewed through the spectroscope. In this case, the spectrum has light at all wavelengths *except* for certain positions where the intensity is diminished. This is known as a **dark line spectrum** or an **absorption spectrum**.

Kirchhoff also noted that the pattern of lines observed in either the emission or absorption spectrum depends on the composition of the light source – in particular, the composition of the hot, thin gas for the emission spectrum or the cooler, thin gas for the absorption spectrum. An optically thin gas of a particular element emits or absorbs at wavelengths that are characteristic of that element; a gas composed of a mixture of elements will exhibit a spectrum that contains a mixture of lines from the various elements.

Line Formation

The fact that absorption lines are observed in the spectra of stars implies that the stars somehow satisfy the geometry of the third law, namely that the continuous spectrum of light from a hot, opaque body must be forced to pass through a cooler, thin gas on its way to the observer. One could reasonably assert that the hot, opaque surface of a star will serve as a convenient source of the continuous spectrum, leaving in doubt only the location of the cooler, thin gas where the lines are formed. Several possibilities come to mind.

The Earth's atmosphere is an optically thin gas, almost certainly apt to be cooler than the surface of a star; and light from any star must pass through our atmosphere on its way to our ground-based telescopes. However, if this is the answer, then our atmosphere should have the *same* effect on *all* stellar spectra; we should be seeing the same absorption lines due to components of the terrestrial atmosphere in the spectrum of every star we observe. And we *could* see them, if we were looking for them, but in most cases astronomers are not interested in studying these terrestrial atmospheric lines – called **telluric lines** – because they are not trying to study the Earth. Such lines are usually ignored or avoided.

One might also imagine the presence of huge clouds of optically thin gas in space. We see such clouds here and there in the sky, illuminated by radiation from nearby stars, but there are certainly cold, dark clouds that are essentially invisible to our gaze. Passage of starlight through these interstellar clouds should produce detectable absorption features in stellar spectra – and it does. However, if this is the main cause of stellar absorption lines, then any two stars in close proximity to each other – such as a binary star – should exhibit nearly identical spectral lines. Furthermore, there should be a strong correlation between the number and intensity of spectral lines and the star's distance – at least for stars near the Galactic plane; nearby stars should have no strong lines. As neither of these predictions is generally observed, there must be another explanation.

Stars, being huge balls of gas, lack distinct boundaries and hard surfaces. Instead, gravity does its best to confine the particles of a gas to a spherical region of space, but there will always be some particles that attain sufficient upward velocities and avoid major collisions long enough that they may exist in the **stellar atmosphere** – a loosely defined region of optically thin gas above the stellar surface. As photons emerge from the surface, they must pass through the stellar atmosphere on their way to us, giving the star a chance to apply its own unique stamp to its spectrum, or to conform to the default spectrum for stars of similar size and temperature.

Just how the star accomplishes this task has not always been obvious. Early stellar models were based on a **photosphere** – literally 'ball of light' – producing the continuous blackbody spectrum, and a **reversing layer** containing metallic vapors that scattered light at discrete wavelengths to make the dark lines. By now the concept of a reversing layer has long since been replaced by one in which the absorption lines are considered to be formed *throughout* the photosphere, but primarily in the higher, cooler layers. Thus, the photosphere produces the entire spectrum, with different regions contributing more heavily to the continuum or to the lines, but there is no longer an attempt to envision distinct layers.

Let us assume that the observer sees into the star to some nominal optical depth – say $\tau_v = 10$. (The attenuation at this value [$(1 - e^{-\tau_v}) = (1 - e^{-10}) \approx 0.99995$] is sufficiently strong for our assertion.) This will be true whether we are looking at a frequency that is in one of the lines or out in the continuum; that is, $\tau_{line} = \tau_{continuum}$. For each region we may write $\tau \approx \kappa\rho x$, where x represents the depth of the atmosphere from which the rays emerge, and the other quantities have their usual meanings. We then have the following condition:

Eq. 5.1 $$\kappa_{line}(\rho\, x)_{line} = \kappa_{continuum}(\rho\, x)_{continuum}$$

Of course the opacity in the line is greater than in the continuum because the line frequency matches a transition frequency for some atomic species while the continuum frequency does not. This leads to the resulting statement about density and depth:

Eq. 5.2 $$(\rho\, x)_{line} < (\rho\, x)_{continuum}$$

Thus, compared to the continuum, the line forms at lower density and/or shallower depth, both of which translate into lines forming high in the atmosphere. Because temperature generally diminishes with altitude, the lines are forming in *cooler* regions, which would then have source functions that are less than those for the continuum. Lower source functions for the line frequencies mean that radiation at these frequencies will be less intense than at the continuum frequencies, and therefore the lines will be darker (as is observed).

Column Density

The darkness of a line is dependent on the opacity at that frequency; it is also a measure of the number of absorbers (of photons of the line frequency) that lie above the stellar surface (where the continuum photons are produced). A useful concept is the **column density n_c**, which is the number of absorbers above each projected unit area of the surface along the line of sight. This can be approximated by the integral of the number density (of the species of interest), from the stellar surface (where $r = R$) to the observer (at $r = \infty$):

Eq. 5.3 $$n_c \approx \int_R^\infty N(r)dr$$

Greater column density will provide a greater number of absorbers, which should produce a stronger (darker, wider) line. Of course, as we have already seen, the line strength is also dependent on the conditions within the atmosphere, such as temperature and pressure (or density), which can determine whether or not a given atom is a potential absorber.

So we have the stellar surface – loosely defined as the radius at which the continuum optical depth approaches infinity – providing a continuous blackbody spectrum characterized by some effective temperature (T_e), with the photosphere (or atmosphere) above this level producing the absorption lines. The structure of the atmosphere is determined by the effective temperature and also by pressure (or density), due to the role the latter plays in ionization. However, because pressure (or density) varies throughout the atmosphere, we need a parameter that we can use to characterize the atmospheric structure in the same way that we use the effective temperature. This parameter is the **surface gravity** of the star: $g = GM/R^2$.

Comparing two stars of the same mass, we find that the star with the larger radius will have a lower surface gravity. This means that the atmosphere of the larger star will not be as tightly bound to the stellar surface, resulting in an extended atmosphere. Such atmospheres require less pressure to counteract the lower gravitational forces. In contrast, a smaller radius will produce a stronger surface gravity, which will pull the atmospheric particles down closer to the surface, resulting in higher pressures. Thus, stars with the same temperature but different surface gravities may have different atmospheres and different spectra.

The value of g for the Sun is given by $\log g \approx 4$; a star with the same mass and a radius 100 times as great would have $\log g \approx 0$. This range of values covers most of the stars, but not all.

Because the spectra of stars are affected by the stars' effective temperatures (via the Boltzmann and Saha equations) and by their surface gravities/pressures (via the Saha equation), we can analyze these spectra to give us information about the temperatures and gravities/ pressures/densities in the stellar atmospheres. Knowledge of opacities allows theoretical predictions of line strengths for different combinations of temperature, pressure, and abundance. These predictions can then be matched to observed spectra to reveal the stellar conditions.

Of course, this has only become feasible with the development of modern computers. Before that, astronomers had to rely on centuries of observational data covering thousands of stars in hopes of finding the keys to stellar behavior. One of the fundamental tools to understanding the similarities and differences among stars was **spectral classification**.

Spectral Classification

Spectral classification was begun in the 19th century as astronomers began examining the spectra of stars and discovering that they were not all the same. It was quite natural to sort the stars into groups according to the appearance of their spectra, assigning letters of the alphabet to distinguish the groups. However, the underlying *causes* of these spectral differences were not immediately obvious to those working in the field (due to lack of suitable models of the atom, etc.), and thus, the alphabetical order of the initial classification schemes did not correspond to any particular physical properties of the stars. Only later, when it became clear that the principal variable affecting stellar spectra was temperature, did our present system of spectral classification emerge.

Harvard System

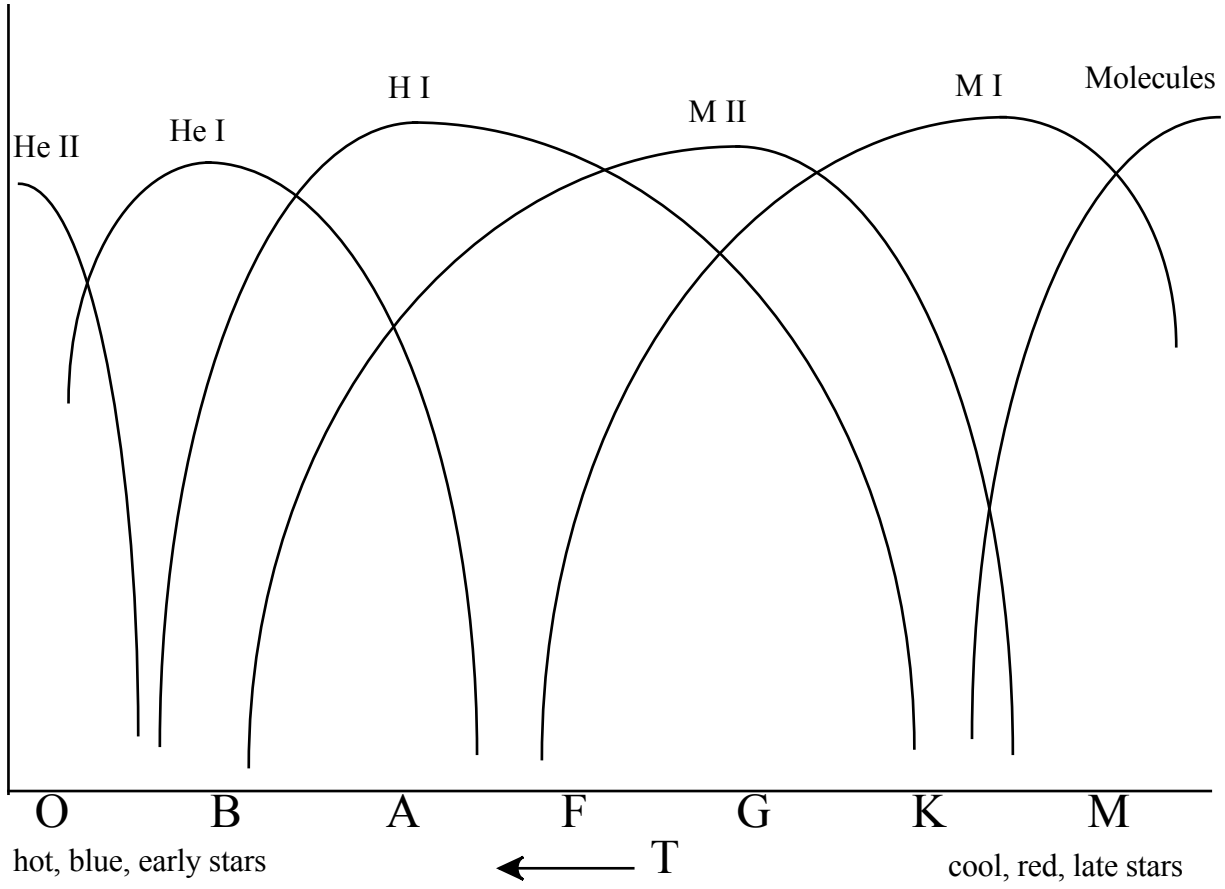
The **Harvard system** utilizes seven basic **spectral types** – O B A F G K M, in order of decreasing temperature – to classify stellar spectra; subclasses 0 through 9 are also used, although not all subclasses are used for each type. Classification is based on the appearance of spectral lines: which ones are present/absent/strong/weak. Table 5.1 indicates the characteristic features of these spectral types, along with approximate temperature ranges.

Figure 5.1 shows the approximate variation in line strengths of different species over the range of spectral types. (These variations were explained at the end of Chapter 4 and depicted in Figure 4.5.) M I and M II refer to neutral and ionized metal species, respectively. Although the metals have a variety of different ionization potentials, they are generally less than the ionization potential for hydrogen, which in turn is lower than that for helium. Thus, metallic species ionize first in the atmospheres of cool stars while helium ionization is only apparent in the hottest stars.

Table 5.1: Characteristics of spectral types

Sp	T_e	Spectral Features
O	> 30000	Strong UV continuum; He II absent; highly ionized atoms; H lines weak.
B	30000 - 9900	He I max at B2; He II vanishes beyond B0; H lines develop in late types.
A	9800 - 7400	H lines max at A0; Ca II increases; weak neutral metal lines appear.
F	7300 - 6000	H lines weaken, but conspicuous; Ca II becomes stronger; metals (ions and neutrals) increase.
G	5900 - 5200	Ca II H & K become strong; Fe and metals strong; H line weak; CH bands strengthen.
K	5100 - 3900	Metallic lines dominate; continuum becomes weak in blue; molecular bands (CN, CH) develop.
M	< 3900	TiO bands dominate; strong neutral metal lines.

Figure 5.1: Relative line strength vs. spectral type

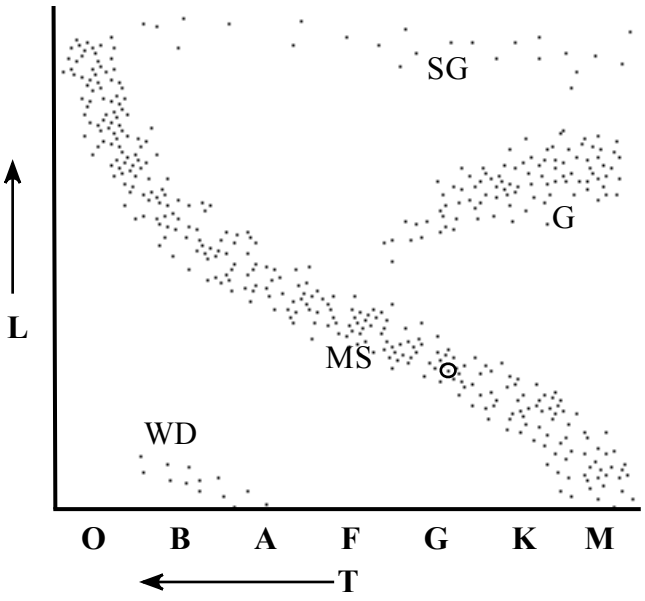


The spectral type and the color of a star are observational quantities that are indicators of the stellar temperature (which is not measured directly). As such they have both been used to characterize stars, most notably in the Hertzsprung-Russell diagram.

The HR diagram was originally developed independently by the American astronomer, Henry Norris Russell and the Danish chemical-engineer-turned-astronomer, Ejnar Hertzsprung. In 1911, Hertzsprung plotted apparent magnitudes vs. color indices for stars in an individual

cluster (which are all at essentially the same distance). He noticed a 'main sequence' of hot stars for both the Pleiades and the Hyades clusters and also a few luminous red stars for the latter cluster. In 1913, Russell used measured stellar parallaxes to determine absolute magnitudes for relatively nearby stars and plotted these magnitudes against Harvard spectral types to obtain his 'Russell diagram'. This also featured a main sequence, but the stars divided into two groups at the cooler spectral types, which Russell termed 'giants' and 'dwarfs'. It soon became evident that the two astronomers were achieving the same result with slightly different techniques; in 1933 their efforts were combined as the Hertzsprung-Russell diagram, a version of which is shown in Figure 5.2. (The small circle marks the Sun's position.)

Figure 5.2: An HR diagram, with the main sequence, giants, supergiants, and white dwarfs indicated



Note that at a given temperature, there can be stars with different luminosities. Russell's 'giants' are still known as giants, and his 'dwarfs' are the cooler main sequence stars – where the division is observed.

There are lesser numbers of supergiants across the top of the diagram. These had distinguished themselves to early classifiers by their extremely sharp, narrow lines, and had been determined to have high luminosities. The white dwarfs exhibit limited spectral features, which are highly broadened, indicating high-pressure atmospheres. Typical surface gravities for white dwarfs are around $\log g \approx 7$.

The size of a star can be calculated if its luminosity and surface temperature are known. From Equation 2.118 we have $F = \sigma T_e^{\,4}$, and the flux is equal to the luminosity divided by the star's surface area:

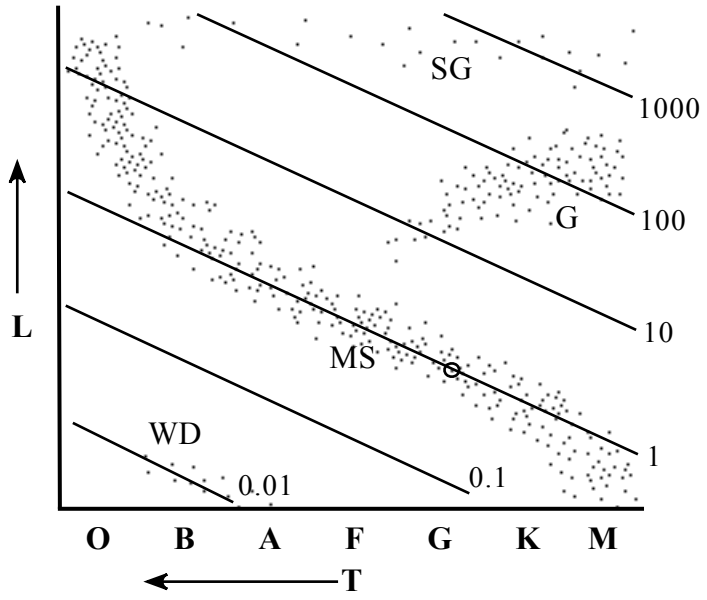
Eq. 5.4 $$F = \frac{L}{4\pi R^2}$$

Combining these gives the desired relation:

Eq. 5.5 $L = 4\pi R^2 \sigma T_e^4$ or $\log L = 2 \log R + 4 \log T_e + \text{constant}$

Stars with the same radius will fall on a straight line across a plot of log L vs log T (an HR diagram), as shown in Figure 5.3.

Figure 5.3: Stars of constant radii, in solar units

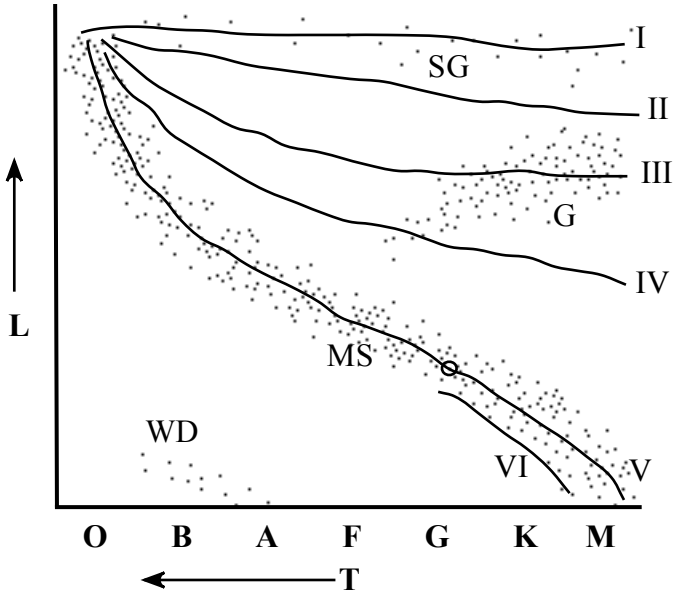


Luminosity Class

In addition to temperature effects there are also variations observed in spectra that are due to pressure (or density) effects. Some of these involve the enhanced ionization produced by low pressure; others are due to the reduced collision rate in atmospheres at low pressure. Collisions between atoms tend to perturb the energy levels of the outer electrons – those usually involved in transitions. If a transition occurs *during* a collision, the photon absorbed will have a frequency slightly different from its normal value, effectively broadening the line. Stars with higher pressure atmospheres will thus exhibit greater **pressure broadening** and be distinguishable from stars with lower pressure atmospheres, which should produce narrower lines.

As described above, the pressure is characterized by the surface gravity, which in turn depends on the size of the star. At a given temperature, larger stars will have lower surface gravity, lower pressure, less broadening, and sharper lines, while smaller stars should have broader, fuzzier lines. Within an individual spectral sub-type, stars may be classified according to the sharpness or fuzziness of their lines, which effectively separates the stars by size, and also by luminosity (as their temperatures will all be similar). The ability to assign a **luminosity class** to a star based on the appearance of its spectral lines has become essential to our understanding of stellar distributions and stellar evolution. The basic luminosity classes used today are designated by Roman numerals, ranging from I (sharpest lines) to V or VI (fuzziest lines), as shown in Figure 5.4.

Figure 5.4: Luminosity classes on the HR diagram



Luminosity class names are related to the established HR diagram groups, as shown in Table 5.2. Luminosity class designations are commonly appended to spectral types, as B3V or F8I. The Sun is a G2V – a main sequence star.

Table 5.2: Luminosity classes

I	supergiants
II	bright giants
III	giants
IV	subgiants
V	main sequence (dwarfs)
VI	subdwarfs

Additional Nomenclature

Additional spectral nomenclature has been developed to designate particular features:

Before the current luminosity classes were introduced, prefixes were used to indicate supergiants (c – based on their fine, narrow lines), giants (g), and main sequence dwarfs (d), with the giants and dwarfs only used for spectral types F through M. Examples are cF8, gM2, and dK5.

Emission lines are relatively rare in stellar spectra, but they do occur in the following cases:

- Wolf-Rayet stars are O stars with emission lines of carbon (type WC) or nitrogen (type WN).
- Some O stars show emission lines of He II and N III (type Of).
- Long-period variable stars (LPVs) show periodic emission lines of H I and other elements, due to the passage of shock waves through the stellar atmosphere (usually type Me).
- Novae (type Q) show emission lines.
- Some B stars show emission lines (type Be).

- Late M dwarfs may be chromospherically active, with prominences – see Kirchhoff's second law – providing a large fraction of the star's luminosity (type dMe).

Additional spectral types have been developed to cover stars of different compositions and temperatures:

- Late giant stars may show ZrO molecular bands (type S) rather than the TiO bands normally found in M stars. Type C stars show bands of different carbon compounds, such as CH, CN, and C_2. These differences are believed to be due to the relative abundances of carbon and oxygen in the stellar atmosphere: M stars have O > C, C stars have O < C, and S stars have O ≈ C. Such differences may be caused by dredge-up of carbon from the core (see Chapter 13).
- In recent years, the cool end of the spectral sequence has been extended to include type L (with $T \approx 1300$ to 2000), type T (with $T \approx 700$ to 1300), and type Y (with $T < 700$). Such 'stars' will emit primarily in the infrared.

Molecular Spectra

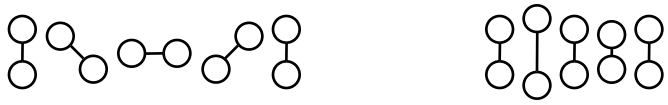
While the atmospheres of stars are far too hot for most molecules to exist, there are a variety of simple molecules that may be found in and around the cooler stars. Most of these will be diatomic molecules, made from two atoms that join together by sharing electrons in flexible, covalent bonds. Such molecules can perform tricks that spherically symmetric atoms can only dream about.

Rotational and Vibrational States

A diatomic molecule may rotate about an axis perpendicular to its bond (see Figure 5.5a), storing energy as rotational kinetic energy. Quantum mechanics prescribes only certain allowable rotation frequencies, which are indicated by the rotational quantum number *J*. (Note: This *J* does not equal *L* + *S*.) The energy associated with rotation is E_{rot}.

A diatomic molecule may also vibrate along the line of the bond (see Figure 5.5b), storing energy as vibrational kinetic energy. Again, quantum mechanics prescribes only certain allowable vibration frequencies, which are indicated by the vibrational quantum number v. The energy associated with vibration is E_{vib}.

Figure 5.5: A rotating diatomic molecule (a) and a vibrating diatomic molecule (b)



a b

In addition, the electrons in a molecule may store electronic energy E_{elec}, just as is possible in an atom. The total energy of a molecule is then the sum of its electronic, vibrational, and rotational energies:

Eq. 5.6 $$E = E_{elec} + E_{vib} + E_{rot}$$

In practice, spectroscopists use *wave numbers* in place of energies, obtained by dividing E by hc. A simplified rotational term $F(J)$ (in cm^{-1}) is as follows:

Eq. 5.7 $$F(J) = \frac{E_{rot}}{hc} = BJ(J+1) \qquad \text{where } B = \frac{h}{8\pi^2 cI}$$

B is the rotational constant, and the value of I it contains is the rotational inertia of the molecule.

Similarly, a simplified vibrational term $G(\text{v})$ is as follows:

Eq. 5.8 $$G(\text{v}) = \frac{E_{vib}}{hc} = \omega\left(\text{v} + \frac{1}{2}\right) \qquad \text{where } \omega = {}^1\!/\!_\lambda$$

However in reality, the rotational model is not independent of the vibrational model, making both terms into series expansions:

Eq. 5.9 $$F = B_\text{v} J(J+1) - D_\text{v} J^2(J+1)^2 + \ldots$$

Eq. 5.10 $$G = \omega_e(\text{v} + {}^1\!/\!_2) - \omega_e x_e(\text{v} + {}^1\!/\!_2)^2 + \omega_e y_e(\text{v} + {}^1\!/\!_2)^3 + \ldots$$

The complete term designation is then of the following form, where T_{elec} is the electronic term:

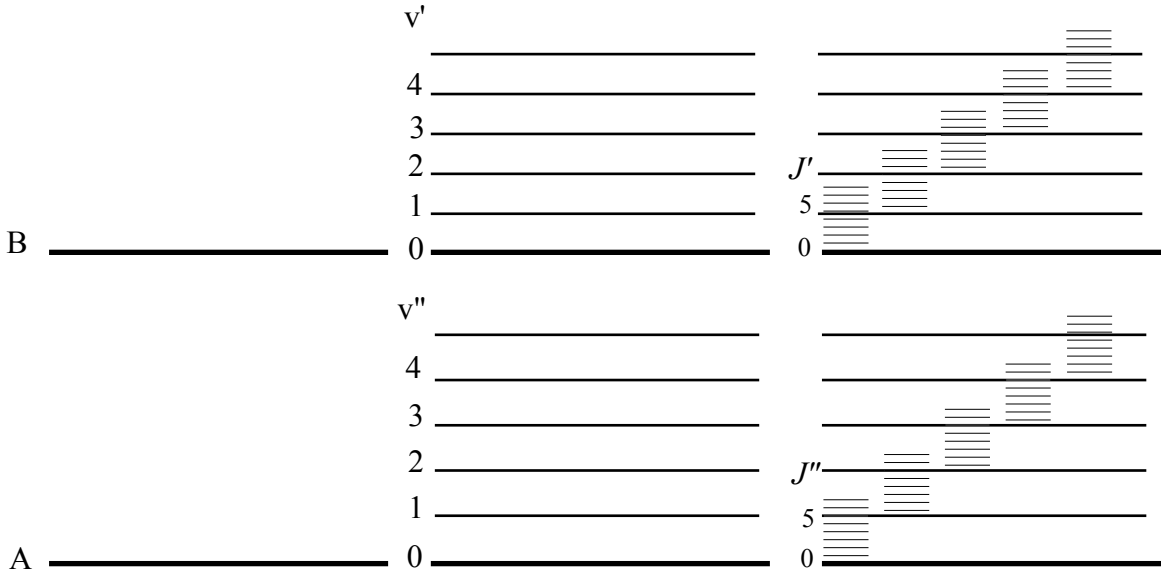
Eq. 5.11 $$T = T_{elec} + G + F$$

In general, the spacing between electronic energy levels will be fairly wide, with vibrational levels more closely spaced and rotational levels *very* closely spaced, as shown in Figure 5.6. Electronic energy level A has several different vibrational levels (v″) associated with it, and each of those has a number of different rotational levels (J''), which are very closely spaced. The upper electronic level (B) is similarly structured. (Note: It is customary to label the upper vibrational and rotational levels in a transition with a single prime and the lower levels with a double prime.)

Transitions between the upper and lower levels may be purely electronic, purely vibrational, or purely rotational in nature, in which case the wavelengths involved will occupy particular regions of the spectrum.

- Purely rotational transitions, with small ΔE are found in the radio and far infrared.
- Purely vibrational transitions, with medium ΔE are found in the near infrared.
- Purely electronic transitions, with large ΔE are found in the visible.

Figure 5.6: Molecular energy levels – electronic, vibrational, and rotational



Band Structure

Alternatively, electronic transitions between two levels (A and B) may also involve changes in the vibrational and/or rotational states of the molecule, subject to quantum mechanical selection rules. In this case the change in energy is given by the sum of the changes in each type of energy:

Eq. 5.12 $$\Delta E = \Delta E_{elec} + \Delta E_{vib} + \Delta E_{rot}$$

In this case, each electronic transition produces not just a single line, but a *set* of lines or **band**. Changes in rotational energy ΔE_{rot} produce a rotational band, with lines spaced typically 1 to 10 Å apart. Changes in vibrational energy ΔE_{vib} then produce *several* rotational bands, spaced on the order of 100 Å apart, as shown in Figure 5.7.

Figure 5.7: Rotational-vibrational bands

Notation

Molecular spectra are thus more complex than atomic spectra, and as might be expected, the notation for molecular spectra is also more complex. But there are a few similarities that will be useful. In general, the notation is about the same, except that for molecular spectra, the Greek alphabet is used, as shown in Table 5.3. As expected, the molecular term notation introduces a few new wrinkles that were not needed for the atomic notation.

Table 5.3: Molecular spectroscopic notation

	Atomic	Molecular
Electron configurations	$\ell = s, p, d, f, g, \ldots$	$\lambda = \sigma, \pi, \delta, \phi, \gamma, \ldots$
Terms	$L = \sum \ell = S, P, D, F, G, \ldots$	$\Lambda = \sum \lambda = \Sigma, \Pi, \Delta, \Phi, \Gamma, \ldots$
Terms of non-equivalent electrons:	$ss \rightarrow {}^1S, {}^3S$	$\sigma\sigma \rightarrow {}^1\Sigma^+, {}^3\Sigma^+$ (+ is for symmetry)
Terms of non-equivalent electrons:	$sp \rightarrow {}^1P, {}^3P$	$\sigma\pi \rightarrow {}^1\Pi, {}^3\Pi_r$ (*r* indicates a regular multiplet)
Terms of non-equivalent electrons:	$pp \rightarrow {}^1S, {}^1P, {}^1D, {}^3S, {}^3P, {}^3D$	$\pi\pi \rightarrow {}^1\Sigma^+, {}^3\Sigma^+, {}^1\Sigma^-, {}^3\Sigma^-, {}^1\Delta, {}^3\Delta_r$

A much more thorough discussion of the spectra of diatomic molecules can be found in Herzberg (1950).

In the next chapter, we will examine the profiles of individual spectral lines to see how they can provide us with further information about the stellar atmosphere.

CHAPTER 6: Line Profiles

In the previous chapter we saw how a star can be assigned to a spectral type, based on the presence or absence, strength or weakness of its absorption lines. We also noted that the *appearance* of its spectral lines can be used to assign a star to a luminosity class. In both cases, we are making use of the fact that for a given transition, a collection of atoms can absorb photons over a small *range* of frequencies, rather than just one (as our earlier discussion of the Bohr model might lead us to believe). The various mechanisms of line broadening and their effects on the line profile are the subject of this chapter; our goal is to be able to deduce information about the stellar atmosphere by examining the profiles of a star's spectral lines.

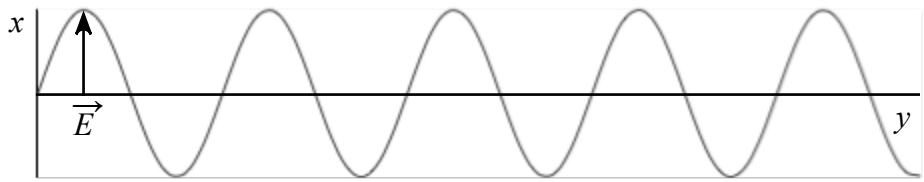
The Natural Line Profile

According to the Bohr model of the atom, electrons may exist only in discrete energy levels. Transitions between two such levels should therefore involve a precise change in the energy of the atom, corresponding to a photon of a certain well-defined frequency: absorption lines should be infinitely sharp. However, this is really not the case. Even in the absence of external broadening mechanisms, the line will be naturally broadened.

The Wave Equation

In order to investigate the interaction of an electromagnetic wave with matter, we begin with the wave itself. Consider an incident electromagnetic wave traveling in the $+y$ direction, with its electric field oscillating in the x direction, as shown in Figure 6.1.

Figure 6.1: The electromagnetic wave (magnetic field not shown)



The function E that gives the magnitude of the oscillating electric field must satisfy the **wave equation**:

Eq. 6.1 $$\frac{\partial^2 E}{\partial t^2} = \mathrm{v}^2 \frac{\partial^2 E}{\partial y^2}$$, where v is the wave velocity.

Any field for which *E* is a function of the quantity ($y \pm \mathrm{v}t$) will satisfy this equation.

The wave velocity v is the speed of light waves in the medium, and this can be found from the following expression:

Eq. 6.2 $$\mathrm{v} = \sqrt{\frac{\varepsilon_o \mu_o}{\varepsilon \mu}} c$$

The quantity ε_o is the electric constant = the **absolute permittivity of free space** (vacuum), while ε is the absolute permittivity of the medium – a measure of the degree to which the medium resists the flow of charge. Some authors use $\varepsilon_r \equiv \varepsilon/\varepsilon_o$ = the **relative permittivit**y = the **dielectric constant**.

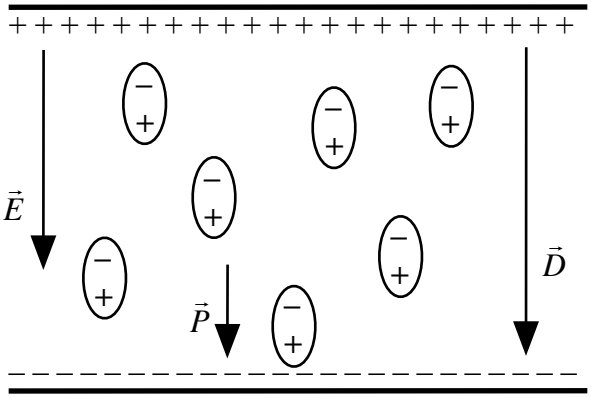
The quantity μ_o is the magnetic constant = the **absolute permeability of free space**, while μ is the permeability of the medium.

The permittivity ε measures the extent of polarization of the medium, while the permeability μ measures the extent of magnetization of the medium. For most gases, $\mu = \mu_o$, and the wave velocity reduces to the following:

Eq. 6.3 $$\mathrm{v} = \sqrt{\frac{\varepsilon_o}{\varepsilon}} c$$

We now recall from electromagnetism that within the medium there are three quantities of interest related to the electric field: $\vec{E}$ is the electric field strength, $\vec{D}$ is the electric displacement, and $\vec{P}$ is the electric polarization. These are illustrated in Figure 6.2, which depicts the electric field inside a parallel-plate capacitor that is filled with a dielectric medium.

Figure 6.2: Electric field vectors within a dielectric medium



The displacement $\vec{D}$ is determined by the free charges on the capacitor plates. The displacement induces molecules in the medium to polarize and align such that their positive ends point in the direction of $\vec{D}$. The alignment of these polarized molecules produces an induced

field ($-\vec{P}$) in the medium that is opposite in direction to $\vec{D}$. The net field is then proportional to the sum of these two fields: $\vec{E} \approx \vec{D} + \left(-\vec{P}\right)$. The actual relations among these three quantities involve the permittivities:

Eq. 6.4 $$\vec{D} = \varepsilon \vec{E} = \varepsilon_o \vec{E} + \vec{P}$$

Eq. 6.5 $$\vec{P} = \left(\varepsilon - \varepsilon_o\right)\vec{E}$$

Now, as noted above, we need an electric field strength that varies with time and with position along the wave in a manner prescribed by the wave equation:

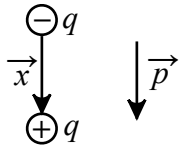
Eq. 6.6 $$E = f\left(y \pm \sqrt{\frac{\varepsilon_o}{\varepsilon}}\, c\, t \right)$$

To accomplish this, we will let the *x*-component of the electric field take on the form of a plane wave traveling in the +*y* direction:

Eq. 6.7 $$E = E_o e^{i\omega\left(t - \sqrt{\varepsilon/\varepsilon_o}\, y/c\right)}$$

We will now let this field interact with the absorbing atoms in the gas, which are modeled as dipole charge oscillators, as shown in Figure 6.3. Here the two charges (+*q* and –*q*) separated by a distance *x* produce a dipole moment $\vec{p} = q\vec{x}$. The electric field induces charge separation in the dipoles, and we can use this interaction to determine a value for $\varepsilon/\varepsilon_o$.

Figure 6.3: The dipole charge oscillator model



If the number of dipoles per unit volume is *N*, and each dipole has a dipole moment $\vec{p} = q\vec{x}$, then the electric polarization produced by the dipole separation is as follows:

Eq. 6.8 $$\vec{P} = N\,\vec{p} = N\,q\,\vec{x}$$

Note that the vectors $\vec{D}, \vec{E}, \vec{P},$ and $\vec{x}$ all point in the same direction. This allows us to remove the vector notation.

Eq. 6.9 $$\vec{D} = \varepsilon \vec{E} \quad \Rightarrow \quad \varepsilon = \frac{D}{E}$$

Eq. 6.10 $$\vec{D} = \varepsilon_o \vec{E} + \vec{P} \quad \Rightarrow \quad \varepsilon_o = \frac{D - P}{E}$$

Combining these equations gives the dielectric constant:

Eq. 6.11 $$\frac{\varepsilon}{\varepsilon_o} = \frac{D}{D-P} = \frac{\varepsilon_o E + P}{\varepsilon_o E} = \frac{E + P/\varepsilon_o}{E} = \frac{E + \dfrac{Nqx}{\varepsilon_o}}{E}$$

Now in cgs units, the constant $4\pi\varepsilon_o$ is equal to 1; therefore, $1/\varepsilon_o = 4\pi$. We will also write the charge q as the electron charge e, as we are dealing with electronic transitions. The dielectric constant can then be written as follows:

Eq. 6.12 $$\frac{\varepsilon}{\varepsilon_o} = \frac{E + 4\pi Nex}{E} = 1 + \frac{4\pi Nex}{E}$$

Damped Harmonic Oscillator

We now consider a single oscillator located at $y = 0$, and driven by the electric field of the wave. At this point the form of the electric field of the wave is simplified (from Equation 6.7):

Eq. 6.13 $$E = E_o e^{i\omega x}$$

We may model this system as a **damped harmonic oscillator**, for which the restoring force is proportional to the displacement, and the damping force is proportional to the velocity of the charge. Writing Newton's second law ($a = \sum F/m$) we find the equation of motion of the charge:

Eq. 6.14 $$\ddot{x} = \frac{e}{m} E_o e^{i\omega t} - \omega_o^2 x - \gamma \dot{x} \qquad \text{(where } m = m_e\text{)}$$

In this equation, the first term on the right represents the driving force (the electric field of the wave), the second term is the restoring force (where ω_o is the natural frequency), and the third term is the damping force (with γ being the **damping constant** – to be discussed later). Rewriting this equation we obtain a second-order linear differential equation:

Eq. 6.15 $$\ddot{x} + \gamma \dot{x} + \omega_o^2 x = \frac{e}{m} E_o e^{i\omega t}$$

We may suppose that the charge will attempt to oscillate in a manner prescribed by the wave (which is driving the oscillation); thus the position of the charge should be given by the following:

Eq. 6.16 $$x = x_o e^{i\omega x}$$

We next find the time derivatives of x:

Eq. 6.17 $$\dot{x} = i\omega x_o e^{i\omega t} = i\omega x \qquad \text{and}$$

Eq. 6.18 $$\ddot{x} = i^2 \omega^2 x = -\omega^2 x$$

Inserting these derivatives into Equation 6.15 gives a linear equation in x:

Eq. 6.19 $$-\omega^2 x + i\gamma\omega x + \omega_o^{\,2} x = (e/m)\, E_o e^{i\omega x}$$

This can easily be solved for x to give the following:

Eq. 6.20 $$x = \frac{e}{m}\frac{E_o e^{i\omega t}}{\omega_o^2 - \omega^2 + i\gamma\omega} = \frac{e}{m}\frac{E}{\omega_o^2 - \omega^2 + i\gamma\omega}$$

This expression can then be substituted into Equation 6.12 to yield the dielectric constant:

Eq. 6.21
$$\begin{aligned}\frac{\varepsilon}{\varepsilon_o} &= 1 + \frac{4\pi N e x}{E} = 1 + \frac{4\pi N e}{E}\left(\frac{e}{m}\frac{E}{\omega_o^2 - \omega^2 + i\gamma\omega}\right) \\ &= 1 + \frac{4\pi N e^2}{m}\left(\frac{1}{\omega_o^2 - \omega^2 + i\gamma\omega}\right) = 1 + \delta\end{aligned}$$

Equation 6.7 requires the square root of this ratio. If we assume that $\varepsilon \approx \varepsilon_o$, then $\delta << 1$ and we can approximate the root by using a binomial expansion:

Eq. 6.22 $$\sqrt{\frac{\varepsilon}{\varepsilon_o}} = \sqrt{1+\delta} \approx 1 + \frac{\delta}{2} = 1 + \frac{2\pi N e^2}{m}\left(\frac{1}{\omega_o^2 - \omega^2 + i\gamma\omega}\right)$$

The complex denominator can be made *real* by using the standard transformation:

Eq. 6.23 $$\frac{1}{a+ib} = \frac{1}{a+ib}\cdot\frac{a-ib}{a-ib} = \frac{a-ib}{a^2+b^2}$$

This yields the following expression:

Eq. 6.24 $$\sqrt{\frac{\varepsilon}{\varepsilon_o}} = \left[1 + \frac{2\pi N e^2}{m}\frac{\left(\omega_o^2 - \omega^2\right)}{\left(\omega_o^2 - \omega^2\right)^2 + \gamma^2\omega^2}\right] - i\left[\frac{2\pi N e^2}{m}\frac{\gamma\omega}{\left(\omega_o^2 - \omega^2\right)^2 + \gamma^2\omega^2}\right] = n - ik$$

The quantity $n - ik$ is the **complex index of refraction**. (Note that n is *not* a quantum number, and k is neither a wave number nor the Boltzmann constant.) Although it was determined by letting $y = 0$, we can apply it to *general* values of y as well:

Eq. 6.25 $$E = E_o e^{i\omega\left(t - \sqrt{\varepsilon/\varepsilon_o}\, y/c\right)} = E_o e^{i\omega(t-(n-ik)y/c)} = E_o e^{i\omega(t-ny/c)} e^{-k\omega y/c}$$

The first exponential factor ($e^{i\omega(t-ny/c)}$) is the plane wave; the second factor ($e^{-k\omega y/c}$) represents the extinction of the wave as it is absorbed by the medium.

The intensity of the radiation is proportional to the *square* of the amplitude (E^*E, where E^* is the complex conjugate). The intensity is then as follows:

Eq. 6.26 $$I = I_o(e^{-i\omega(t-ny/c)}e^{-k\omega y/c})(e^{i\omega(t-ny/c)}e^{-k\omega y/c}) = I_o e^{-2k\omega y/c} = I_o e^{-\tau} = I_o e^{-\kappa\rho y}$$

And we can use this result to obtain an expression for the absorption coefficient:

Eq. 6.27 $$\kappa_\omega \rho = \frac{2k\omega}{c} = \frac{4\pi N e^2}{mc}\frac{\gamma\omega^2}{\left(\omega_o^2 - \omega^2\right)^2 + \gamma^2\omega^2}$$

This is a sharply peaked function of ω that is non-zero only when $\omega \approx \omega_o$ – that is, when the photon frequency ω matches the transition frequency ω_o. When these two frequencies are close, we can make the following simplifying approximation:

Eq. 6.28 $$\omega_o^2 - \omega^2 = (\omega_o + \omega)(\omega_o - \omega) \approx 2\omega(\omega_o - \omega) = 2\omega\,\Delta\omega$$

Then the absorption coefficient in the region of the line is as follows:

Eq. 6.29 $$\kappa_\omega \rho \approx \frac{4\pi N e^2}{mc} \frac{\gamma\omega^2}{4\omega^2(\Delta\omega)^2 + \gamma^2\omega^2} = N\frac{\pi e^2}{mc}\frac{\gamma}{(\Delta\omega)^2 + (\gamma/2)^2}$$

Dispersion Profile

Figure 6.4 shows the typical form of this function; in general, the curve is sharply peaked at $\Delta\omega = 0$. It goes by several names, including the **dispersion profile**, the **damping profile**, the **Lorentzian profile**, the **Cauchy curve**, and the **witch of Agnesi**.

Figure 6.4: A dispersion profile

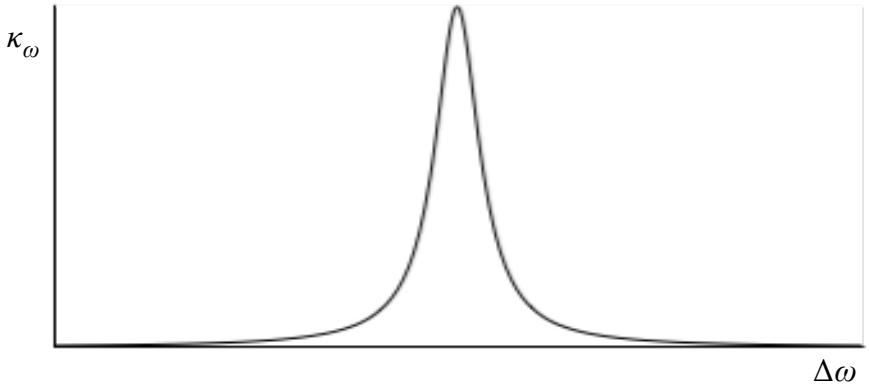
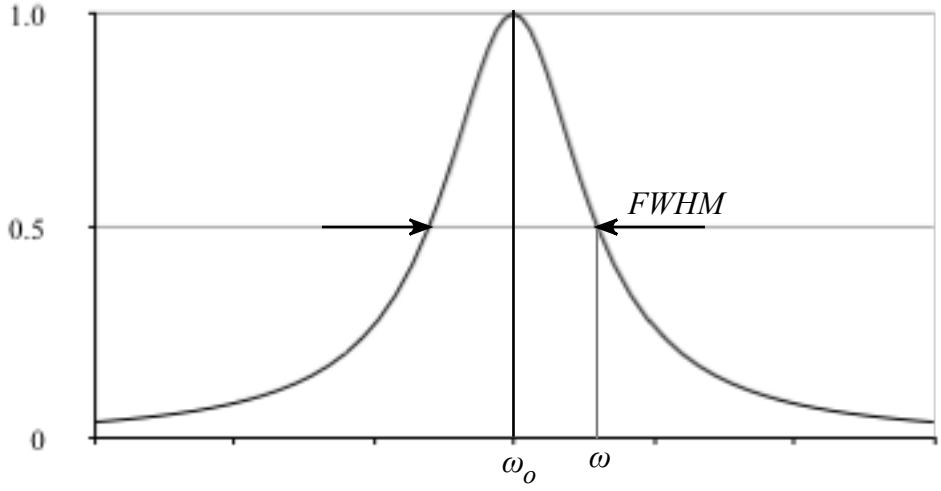


Figure 6.5: The halfwidth (FWHM) of a dispersion profile



A useful concept in describing line profiles is the halfwidth, or – more precisely – the **full width at half maximum (FWHM)**. This is found by measuring the width of the profile at a frequency ω where the value of the profile is one half the maximum value (which of course is located at ω_o). The value of the halfwidth is then given by FWHM = $2(\omega - \omega_o) = 2\Delta\omega$. The halfwidth is illustrated in Figure 6.5.

Absorption Cross Sections

The quantity given as the absorption coefficient in Equation 6.29 is not quite what we want, as it gives the mass absorption coefficient κ as a function of the *angular* frequency ω, rather than the frequency ν. We can modify this relation by writing the dispersion profile in terms of cross sections, using the following simple transformation:

Eq. 6.30 $\kappa\rho = N\sigma \quad \Rightarrow \quad \sigma = \kappa\rho/N$

In terms of angular frequency ω, we have the following cross section:

Eq. 6.31 $$\sigma_\omega = \frac{\kappa\rho}{N} = \frac{\pi e^2}{mc}\frac{\gamma}{(\Delta\omega)^2 + (\gamma/2)^2} = \frac{2\pi e^2}{mc}\frac{\gamma/2}{(\Delta\omega)^2 + (\gamma/2)^2}$$

This cross section has a halfwidth given by $FWHM = \gamma = 2\,\Delta\omega$, where $\Delta\omega = \omega - \omega_o$.

Similarly, employing the relations $\omega = 2\pi\nu$ and $\lambda = c/\nu$, we may obtain cross sections in terms of frequency and wavelength, along with their accompanying half-widths:

Eq. 6.32 $$\sigma_\nu = \frac{e^2}{mc}\frac{\gamma/4\pi}{(\Delta\nu)^2 + (\gamma/4\pi)^2} \qquad FWHM = \gamma/2\pi = 2\Delta\nu \qquad \Delta\nu = \nu - \nu_o .$$

Eq. 6.33 $$\sigma_\lambda = \frac{e^2}{mc}\frac{\lambda^2}{c}\frac{\gamma\lambda^2/4\pi c}{(\Delta\lambda)^2 + (\gamma\lambda^2/4\pi c)^2} \qquad FWHM = \gamma\,\lambda^2/2\pi c = 2\Delta\lambda \qquad \Delta\lambda = \lambda - \lambda_o .$$

Now to find the *total* absorption of a line, we could integrate the line profile over all frequencies (ignoring the fact that we obtained the line profile by assuming the frequency was in the vicinity of the transition frequency). The total absorption of a line is then given by the following:

Eq. 6.34 $$\int_0^\infty \sigma_\nu d\nu = \int_{-\infty}^{\infty} \sigma_\nu d(\Delta\nu) = a\int_{-\infty}^{\infty} \frac{b}{x^2 + b^2} dx \qquad \text{where } a = \frac{e^2}{mc} \text{ and } b = \frac{\gamma}{4\pi}$$

This integral can easily be solved:

Eq. 6.35 $$a\int_{-\infty}^{\infty} \frac{b}{x^2 + b^2} dx = ab\left(\frac{1}{b}\right)\tan^{-1}\frac{x}{b}\Big|_{-\infty}^{\infty} = a\left(\frac{\pi}{2} - \frac{-\pi}{2}\right) = a\pi = \frac{\pi e^2}{mc}$$

Evaluation of this expression gives a value of 0.02654 ergs/(sec-ster-atom) – the energy absorbed by the total line from a unit I_ν beam. Alternatively, we can find the total energy taken from an I_λ beam:

Eq. 6.36 $$\int_0^\infty \sigma_\lambda d\lambda = \int_{-\infty}^{\infty} \sigma_\lambda d(\Delta\lambda) = \frac{\pi e^2 \lambda^2}{mc^2}$$

Oscillator Strength

These are *theoretical* values. Actual measurements show these amounts to be too *large*. Therefore we will now introduce a factor f (from a quantum mechanical treatment of the problem) such that it scales our theoretical value to the correct result:

Eq. 6.37 $$\int_0^\infty \sigma_\nu d\nu = \frac{\pi e^2}{mc} f$$

This factor is known as the **Ladenburg *f* value** or the **oscillator strength**. It is equivalent to the fractional number of electron oscillators per atom for a given transition.

The oscillator strength is different for each level and is related to the Einstein B-coefficient B_{12} (from Chapter 4) because we must have $\int_0^\infty \sigma_\nu d\nu = \frac{h\nu}{4\pi} B_{12}$. This requires the following:

Eq. 6.38 $$f = \frac{mc}{4\pi^2 e^2} h\nu B_{12}$$

Equations 4.19 and 4.20 give $A_{21} = \frac{2h\nu^3}{c^2}\frac{g_1}{g_2} B_{12}$, which produces the following relations:

Eq. 6.39 $$f = \frac{mc^3}{8\pi^2 e^2 \nu^2}\frac{g_2}{g_1} A_{21} = \frac{mc\lambda^2}{8\pi^2 e^2}\frac{g_2}{g_1} A_{21}$$

This is an oscillator strength for absorption (f_{abs}); We can also define an oscillator strength for emission (f_{em}) by utilizing the statistical weights of the upper and lower levels:

Eq. 6.40 $$g_u f_{em} = g_\ell f_{abs}$$

Allen (1973) and Cox (2000) give tables of gf, which can be used for either product. Oscillator strengths can also be determined empirically in the lab. For hydrogen, Kramers' formula gives the following:

Eq. 6.41 $$f = \frac{2^5}{3\sqrt{3}\pi}\frac{g_{bb}}{n_\ell^5 n_u^3}\left[\frac{1}{n_\ell^2} - \frac{1}{n_u^2}\right]^{-3}$$

Here n_ℓ and n_u are quantum numbers for the lower and upper levels, and g_{bb} is the bound-bound Gaunt factor, which has a value on the order of 1. Some calculated oscillator strengths for Balmer lines are given in Table 6.1.

Table 6.1: Oscillator strengths for Balmer lines (Gray 1976, Bowers & Deeming 1984)

Line	f
Hα	0.640742
Hβ	0.119321
Hγ	0.044670
Hδ	0.022093
Hε	0.012704
Hζ	0.008036

Damping Constant for Radiation Damping

We now turn our attention to determination of a value for the **damping constant for natural broadening** (radiation damping), which was introduced in Equation 6.14. As our subsequent calculations have shown, the damping constant is proportional to the natural broadening halfwidth and is thus of some interest.

We can obtain a *classical* value for the damping constant γ from the rate at which an accelerated charge loses energy (W):

Eq. 6.42 $$\frac{dW}{dt} = -\gamma W \quad \Rightarrow \quad W = W_o e^{-\gamma t}$$

This results in the following expressions for γ, in terms of ω, ν, and λ:

Eq. 6.43 $$\gamma = \frac{2}{3}\frac{e^2}{mc^3}\omega^2 = \frac{8\pi^2}{3}\frac{e^2}{mc^3}\nu^2 = \frac{8\pi^2}{3}\frac{e^2}{mc}\frac{1}{\lambda^2}$$

A numerical value for γ is 2.47e–22 ν^2 s^{-1}. This results in a halfwidth (in λ form) as follows:

Eq. 6.44 $$FWHM = \frac{\gamma\lambda^2}{2\pi c} = \frac{4\pi}{3}\frac{e^2}{mc^2} = 1.18\times10^{-4} \text{ Å for all lines}$$

This classical result is *smaller* than what is observed by an order of magnitude, indicating a quantum mechanical treatment is needed. We begin by assuming the energy is quantized (because energy is absorbed and emitted in photons):

Eq. 6.45 $W = N_u h\nu$ where N_u is the upper level population

Then the rate at which energy is emitted is as follows:

Eq. 6.46 $$\frac{dW}{dt} = -\gamma W \quad \Rightarrow \quad \frac{dN_u}{dt} = -\gamma N_u$$

Now the rate at which electrons leave the upper level must be equal to the rate at which electrons enter all the various lower levels:

Eq. 6.47 $$\frac{dN_u}{dt} = \sum_\ell \frac{dN_{u\ell}}{dt} = -N_u \sum_\ell A_{u\ell}$$

Here $N_{u\ell}$ is the number of downward transitions and $A_{u\ell}$ is the Einstein A-coefficient.

Comparing these last two equations, we find an expression for the damping constant:

Eq. 6.48 $$\gamma_u = \sum_{\ell} A_{u\ell}$$

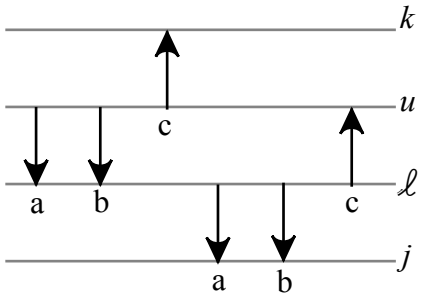
Then the lifetime of the excited state is the inverse of this value:

Eq. 6.49 $$\Delta t = \frac{1}{\gamma_u} = \frac{1}{\sum_{\ell} A_{u\ell}}$$

In general, we should also allow for stimulated emission (to level ℓ) and absorption to higher levels (k):

Eq. 6.50 $$\gamma_u = \sum_{\ell<u} A_{u\ell} + \sum_{\ell<u} I_\nu B_{u\ell} + \sum_{k>u} I_\nu B_{uk}$$

Figure 6.6: Transitions into and out of the upper (u) and lower (ℓ) levels of interest, via spontaneous emission (a), stimulated emission (b), and absorption (c)



We can write a similar expression for the damping constant for the lower level ℓ, where the levels are as shown in Figure 6.6. These are needed for *strong* radiation fields.

Eq. 6.51 $$\gamma_\ell = \sum_{j<\ell} A_{\ell j} + \sum_{j<\ell} I_\nu B_{\ell j} + \sum_{u>\ell} I_\nu B_{\ell u}$$

Thus *each* energy *level* is broadened; the broadening of the *line* is then a combination of the broadening for *both* levels because a given transition may combine a range of different lower levels with a range of different upper levels.

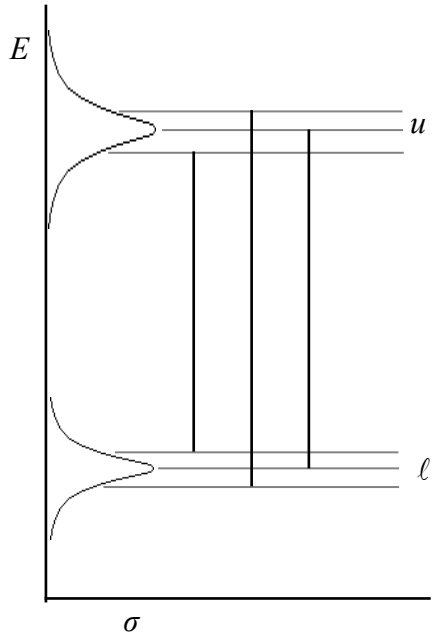
Eq. 6.52 $$\gamma_{lin\ vbge} = \gamma_u + \gamma_\ell$$

Figure 6.7 shows how the broadening of each energy level contributes to the overall broadening of the line, by allowing photons of a range of energies to participate in the transition.

Note that *all* levels must be broadened, except for the ground state, which may be infinitely sharp. This is due to the uncertainty principle, which says that $\Delta x \Delta p \geq \hbar$. This relation can be manipulated to yield the alternative form:

Eq. 6.53 $$\Delta x \Delta p = (\Delta x \Delta p / \Delta t)\Delta t = (\Delta x \Delta F)\Delta t = \Delta E \Delta t \geq \hbar$$

Figure 6.7: Broadening of the line produced by the two broadened levels



Because an atom may remain in the ground state for an extremely long time, the uncertainty in the energy of the ground state ($\Delta E \geq \hbar/\Delta t$) may be quite small. Excited states – which have much shorter lifetimes – must have greater uncertainties in their energies.

This type of broadening is called **radiation damping**, and the value of γ depends on *all* possible transitions in and out of *both* levels involved in the transition of interest. Radiation damping depends only on atomic properties and is thus independent of the atom's environment. This means that if other (external) broadening mechanisms are absent or minimal, the line profile will convey *little* information about the state of the atmosphere. Normally, this is not the case however, as there are plenty of other ways in which a line may be broadened.

Pressure/Collisional Broadening

Pressure broadening, which was mentioned in the previous chapter in conjunction with the determination of luminosity class, is caused by collisional interactions between absorbing atoms and other atoms, ions, electrons, etc. These collisions perturb the atomic energy levels and thus alter the energy of a transition; upper (outer) levels are generally affected more.

The general form of pressure broadening is given by $\Delta\nu = C_n/R^n$, where R is the separation between the absorbing atom and the perturber, and C_n is an interaction constant, which is different for each transition and generally unknown. Table 6.2 lists several types of pressure broadening.

These all result in dispersion profiles with widths γ_n. The total dispersion profile then has a width $\gamma = \sum \gamma_n$, where the sum is over the halfwidths resulting from all of the different mechanisms that produce dispersion profiles.

Table 6.2: Types of pressure broadening

n	Type	Lines Affected	Perturber
2	Linear Stark	hydrogen	p^+, e^-
4	Quadratic Stark	most lines, especially in hot stars	e^-
6	van der Waals	most lines, especially in cool stars	H

Thermal/Doppler Broadening

A completely different type of line profile results from a mechanism that plays a very important role in stellar spectra: the motions of atoms within a gas. Gas particles move randomly, with different directions and different speeds. The distribution of particle speeds is characterized by the temperature of the gas, which is a measure of the *average* kinetic energy of a particle.

Motions are important because relative motion between the source and the observer produces a Doppler shift – a difference between the frequency of light absorbed by the source and the frequency of the absorption line measured by the observer. Along the observer's line of sight through the gas, some atoms will be moving away from and others toward the observer, resulting in Doppler shifts to frequencies both above and below the transition frequency. This is known as **thermal broadening** or **Doppler broadening**. Only the component of a particle's velocity that is along the line of sight – the **radial velocity** – will contribute to Doppler broadening; tangential components may be ignored.

Doppler Profile

We wish to determine the line profile produced by Doppler broadening for a transition frequency ν_o (the line center in the rest frame) and a radial velocity v_r. The Doppler shift is proportional to the radial velocity:

Eq. 6.54 $$\Delta\nu = \nu - \nu_o = \nu_o \left(\mathrm{v}_r/c\right)$$

If the gas as a whole is at rest with respect to the observer, then the average radial velocity will be zero as there will be equal numbers of atoms moving towards and away from the observer. We are interested in the fraction of atoms $N(\mathrm{v}_r)d\mathrm{v}_r/N$ that have a radial velocity in the range $\mathrm{v}_r \rightarrow \mathrm{v}_r + d\mathrm{v}_r$ and thus have kinetic energy $E = {}^1/_2\, m\mathrm{v}_r^2$ (where m is the mass of an atom); this fraction can be found from Boltzmann's equation:

Eq. 6.55 $$\frac{N(\mathrm{v}_r)d\mathrm{v}_r}{N} = Ce^{-E/kT}d\mathrm{v}_r = Ce^{-m\mathrm{v}_r^2/2kT}d\mathrm{v}_r$$

(No statistical weights appear in this equation because it is assumed that in a normal gas, there are sufficient states available such that particles are not limited by statistical weights.) The constant C is a normalization constant, which we may ignore for the present.

This expression can be simplified further by noting that the most probable speed (the peak speed) in such an ideal gas is given by $\mathrm{v}' = \sqrt{2kT/m}$:

Eq. 6.56 $$\frac{N(\mathrm{v}_r)d\mathrm{v}_r}{N} = Ce^{-\left(\frac{\mathrm{v}_r}{\mathrm{v}'}\right)^2} d\mathrm{v}_r$$

We now use Equation 6.54 to write the radial velocity in terms of frequency:

Eq. 6.57 $$\mathrm{v}_r = \frac{c\Delta\nu}{\nu_o} \qquad \text{and} \qquad d\mathrm{v}_r = \frac{cd(\Delta\nu)}{\nu_o} = \frac{cd\nu}{\nu_o}$$

And we also define a Doppler width $\Delta\nu_D$ in terms of the peak speed:

Eq. 6.58 $$\Delta\nu_D = \nu_o \frac{\mathrm{v}'}{c} = \frac{\nu_o}{c}\sqrt{\frac{2kT}{m}}$$

Then the strength of the absorption in the frequency range $\nu \rightarrow \nu + d\nu$ is proportional to $e^{-(\Delta\nu/\Delta\nu_D)^2}d\nu$ (from Equation 6.56). The normalized frequency distribution function for Doppler broadening is then a Gaussian:

Eq. 6.59 $$\phi(\nu) = \frac{1}{\Delta\nu_D\sqrt{\pi}} e^{-(\Delta\nu/\Delta\nu_D)^2}$$

The normalization constant was obtained by setting $\int \phi(\nu)\, d\nu = 1$, and assuming $\Delta\nu_D << \nu_o$. Because the Doppler width increases with temperature, this profile will also be broader for higher temperatures.

Voigt Function

We now have two different line profiles resulting from different broadening mechanisms, and they both act in combination to broaden the line. The integrated cross section is given by Equation 6.37 as $\pi e^2 f/mc$; the differential cross section for a particular profile is then this value multiplied by the normalized distribution function for the profile. Thus we have the dispersion cross section, where the damping constant is the sum of the γ values for all mechanisms:

Eq. 6.60 $$\sigma_\nu(\Delta\nu) = \frac{\pi e^2}{mc} f \cdot \phi_{disp}(\Delta\nu) = \frac{\pi e^2}{mc} f \left[\frac{1}{\pi}\frac{\gamma/4\pi}{(\Delta\nu)^2 + \left(\gamma/4\pi\right)^2}\right]$$

And similarly, we have the corresponding Doppler cross section:

Eq. 6.61 $$\sigma_\nu(\Delta\nu) = \frac{\pi e^2}{mc} f \cdot \phi_{Dopp}(\Delta\nu) = \frac{\pi e^2}{mc} f \left[\frac{1}{\Delta\nu_D\sqrt{\pi}} e^{-(\Delta\nu/\Delta\nu_D)^2}\right]$$

The two are combined using a convolution.

Convolution

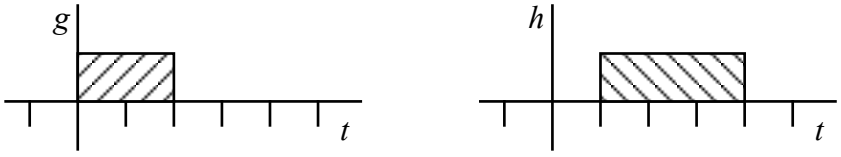
The **convolution (∗)** of two functions is defined as an integral:

Eq. 6.62 $$g(\tau) * h(\tau) \equiv \int_{-\infty}^{\infty} g(t)\cdot h(\tau - t)\, dt$$

We may visualize this process as reflecting the h function across the vertical axis, sliding it past the g function (by varying τ), and measuring the overlap area as a function of the offset τ. Only in regions where *both* functions are non-zero will the product make a contribution to the integral. A simple example will illustrate the basic features.

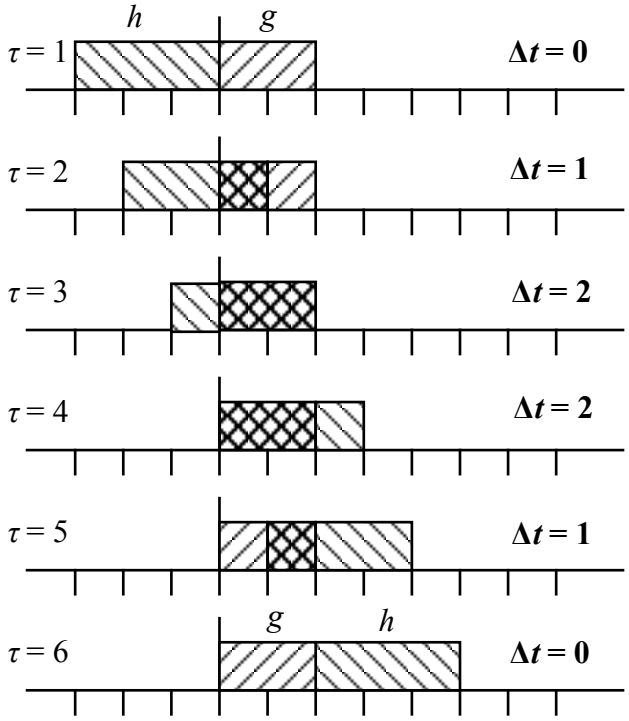
Let $g(t) = \begin{cases} 1 \text{ for } t = 0 \rightarrow 2 \\ 0 \text{ elsewhere} \end{cases}$ and let $h(t) = \begin{cases} 1 \text{ for } t = 1 \rightarrow 4 \\ 0 \text{ elsewhere} \end{cases}$ (See Figure 6.8.)

Figure 6.8: Convolution example functions g(t) and h(t)



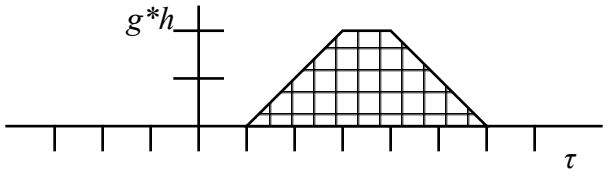
For the convolution integral, we need only consider regions where $g(t) \neq 0$ (true for the interval $t = 0 \rightarrow 2$) *and* $h(\tau - t) \neq 0$ (true for $\tau - t = 1 \rightarrow 4$). These conditions give two inequalities ($1 \leq \tau - t \leq 4$ and $0 \leq t \leq 2$) that can be solved to yield the values of interest for τ : $1 \leq \tau \leq 6$. On this interval, $g = h = 1$, and $g*h = \int dt$. At each offset value τ, we determine the width of the overlap Δt, which gives us the integral $\int dt$ and thus, the convolution.

Figure 6.9: Calculation of the convolution



The result of the convolution is presented graphically in Figure 6.10; for this simple example, it is essentially a plot of Δt vs. τ.

Figure 6.10: Result of the convolution integral for the example



It should be noted that the convolution profile in this case does not closely resemble *either* of the original profiles that combined to form it. This will generally be the case. It should also be noted that convolutions abound in real life – at least in the real lives of *astronomers*. For example, passing starlight through a slit to form a spectrum involves a convolution of the line profile with the profile of the slit, resulting in a modified line profile. Similarly, observing an extended radio source with a radio telescope produces a convolution of the source profile with the antenna beam pattern. In each case, the profile we observe is dependent on both the source and the instrument used to observe it.

For our particular problem, we need to combine a dispersion profile with a Doppler profile:

$$\sigma_\nu(\Delta\nu) = \frac{\pi e^2}{mc} f \cdot \phi_{disp}(\Delta\nu) * \phi_{Dopp}(\Delta\nu)$$

Eq. 6.63

$$= \frac{\pi e^2}{mc} f \left[\frac{1}{\pi} \frac{\gamma/4\pi}{(\Delta\nu)^2 + (\gamma/4\pi)^2} \right] * \left[\frac{1}{\Delta\nu_D \sqrt{\pi}} e^{-\left(\frac{\Delta\nu}{\Delta\nu_D}\right)^2} \right]$$

The convolution in this expression is called the **Voigt function**:

$$V(\Delta\nu, \Delta\nu_D, \gamma) \equiv \phi_{disp}(\Delta\nu) * \phi_{Dopp}(\Delta\nu)$$

Eq. 6.64

$$= \left[\frac{1}{\pi} \frac{\gamma/4\pi}{(\Delta\nu)^2 + (\gamma/4\pi)^2} \right] * \left[\frac{1}{\Delta\nu_D \sqrt{\pi}} e^{-\left(\frac{\Delta\nu}{\Delta\nu_D}\right)^2} \right]$$

$$= \int_{-\infty}^{\infty} \left(\frac{1}{\pi} \frac{\gamma/4\pi}{(\Delta\nu - \Delta\nu')^2 + (\gamma/4\pi)^2} \right) \left(\frac{1}{\Delta\nu_D \sqrt{\pi}} e^{-\left(\frac{\Delta\nu'}{\Delta\nu_D}\right)^2} \right) d(\Delta\nu')$$

Or we may simplify the Voigt function by defining $u \equiv \frac{\Delta\nu}{\Delta\nu_D}$ and $a \equiv \frac{\gamma/4\pi}{\Delta\nu_D}$:

Eq. 6.65 $$V(u,a) = \frac{1}{\Delta\nu_D \sqrt{\pi}} \frac{a}{\pi} \int_{-\infty}^{\infty} \frac{e^{-u'^2}}{(u-u')^2 + a^2} du'$$

This function is tabulated in Gray (1976) and other sources. The cross section is then as follows:

Eq. 6.66 $$\sigma_\nu = \frac{\pi e^2}{mc} f V(u,a)$$

Hjerting Function

The Voigt function can also be defined in terms of the **Hjerting function *H(u,a)***:

Eq. 6.67 $$V(u,a) = \frac{1}{\Delta\nu_D\sqrt{\pi}} H(u,a)$$

Eq. 6.68 $$H(u,a) = \frac{a}{\pi}\int_{-\infty}^{\infty} \frac{e^{-u'^2}}{(u-u')^2 + a^2} du'$$

The Hjerting function is tabulated in polynomial form in Gray (1976) and Allen (1973):

Eq. 6.69 $$H(u,a) = H_0(u) + aH_1(u) + a^2H_2(u) + a^3H_3(u) + ...$$

The damping parameter $a \equiv \frac{\gamma/4\pi}{\Delta\nu_D}$ is typically small, ≈ 0.1, meaning that the Doppler broadening usually dominates. The two profiles affect different parts of the line, with the line core shaped by Doppler broadening and the wings shaped by damping.

Microturbulence

Doppler broadening is caused by random atomic (thermal) motions, but the gas itself may acquire a *bulk* velocity that can also contribute to the line broadening. Along the line of sight through the atmosphere there may be volumes of gas with different radial velocities. These will effectively change the cross section in the line and expand the photons and atoms that can participate in the transition. This is called **microturbulence**, which is a form of Doppler broadening.

Microturbulence can be accounted for by introducing a microturbulent velocity v_t, which further broadens the Doppler core:

Eq. 6.70 $$\mathrm{v}' = \sqrt{\frac{2kT}{m}} \quad \Rightarrow \quad \sqrt{\frac{2kT}{m} + \mathrm{v}_t^2}$$

Macroturbulence

Microturbulence involves motions on a *small scale* compared to the line-forming region. Motions on a *large scale* compared to the line-forming region produce **macroturbulence**, which requires a different approach.

In this case, different areas of the stellar disk or different layers in the atmosphere may have different radial velocities, shifting the profile for that region to the blue or the red. The observed line profile for the star is then the sum of these Doppler-shifted profiles.

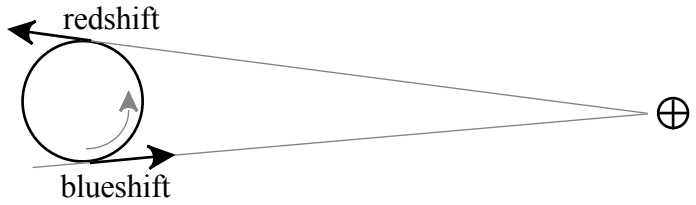
An example of macroturbulence can be found in the atmospheres of long-period variable stars (LPVs), which vary in brightness on time scales of about a year. Shock waves emerge from the photosphere and propagate outwards through the atmosphere, imparting an outward radial velocity to each layer, which changes gradually to an inward velocity as the layer falls back on the star. The line-forming region of the atmosphere thus consists of numerous layers with different radial velocities; which layer dominates the line profile depends on the local

atmospheric conditions, which in turn vary with time as the shock waves move outwards. Absorption lines in an LPV spectrum are thus observed to vary their radial velocity over the period of the star.

Rotational Broadening

Many stars rotate, and this motion can produce radial velocity components and their resulting Doppler shifts for different parts of the star, as shown in Figure 6.11.

Figure 6.11: Doppler shifts produced by stellar rotation

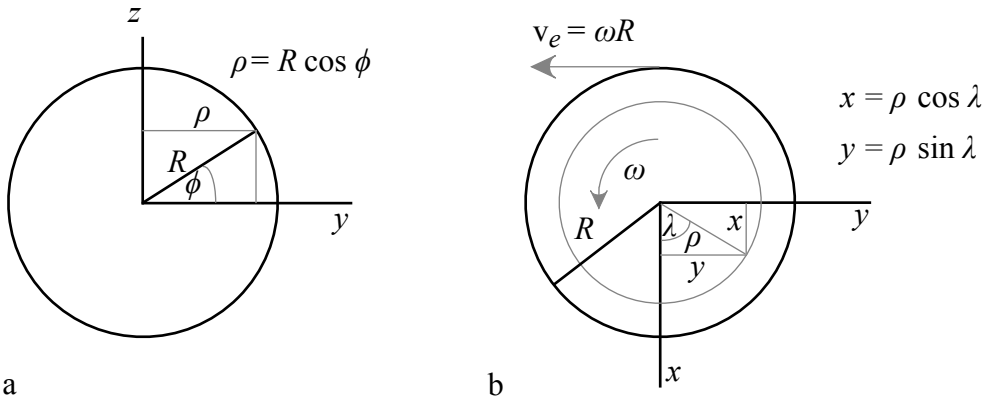


For most stars the disk is not resolved, and photons from all parts of the disk are observed in the same beam. The line profile is then the sum of contributions from the visible hemisphere of the star. As redshifts and blueshifts should be equally probable across the disk, the line will be effectively broadened without being displaced.

To determine the form of the rotational profile, we must establish a coordinate system on the star. For simplicity we will assume the angle of inclination i – the angle between the rotational axis and the line of sight to the observer – to be 90°. Then we select the x-axis as the line of sight and the z-axis as the rotational axis, which places the star's disk in the y-z plane and the observer in the equatorial (x-y) plane.

For an arbitrary point on the surface of the star we may define a latitude ϕ (the angle above the equatorial plane) and a longitude λ (the angle from the central meridian), as shown in Figure 6.12.

Figure 6.12: Coordinate system for rotational broadening, as seen (a) by the observer and (b) from above the pole of the rotating star



The x- and y-coordinates of a general point on the star's surface (λ, ϕ) are as follows:

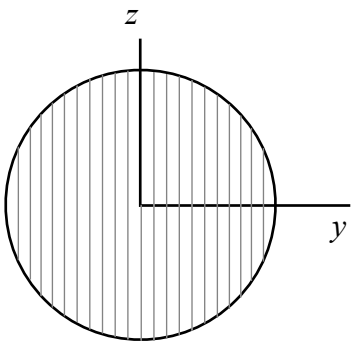
Eq. 6.71 $x = \rho \cos\lambda = R \cos\phi \cos\lambda$

Eq. 6.72 $y = \rho \sin\lambda = R \cos\phi \sin\lambda$

The radial velocity of a point on the star's equator is the *x*-component of the equatorial velocity: $v_r = v_e \sin\lambda = \omega R \sin\lambda$. For a general point ($\lambda$, ϕ), the radial velocity is proportional to the *y*-coordinate:

Eq. 6.73 $v_r = \omega\rho \sin\lambda = \omega R \cos\phi \sin\lambda = \omega y = v_e y/R$

Figure 6.13: The stellar disk, divided into strips of constant radial velocity



We may divide the visible disk of the star into strips of constant *y*, each with radial velocity $v_r = v_e y/R$ and a weight proportional to the length ($= 2z = 2\sqrt{1-y^2}$). We may let $R = 1$ and recall that $v_r/c = \Delta\nu/\nu$, which allows the *y*-coordinate to be written in terms of the Doppler shift:

Eq. 6.74 $$y = \frac{v_r}{v_e} = \frac{c\Delta\nu}{v_e\nu}$$

The rotational broadening profile can then be written as follows:

Eq. 6.75 $$\phi_{rot}(\Delta\nu) \approx 2\sqrt{1-y^2} = 2\sqrt{1-\left(\frac{c\Delta\nu}{v_e\nu}\right)^2}$$

The above was derived for an inclination of 90°; if this is not the case, then the equatorial velocity must be replaced by $v_e \sin i$, giving a modified profile:

Eq. 6.76 $$\phi_{rot}(\Delta\nu) \approx 2\sqrt{1-\left(\frac{c\Delta\nu}{v_e\nu \sin i}\right)^2}$$

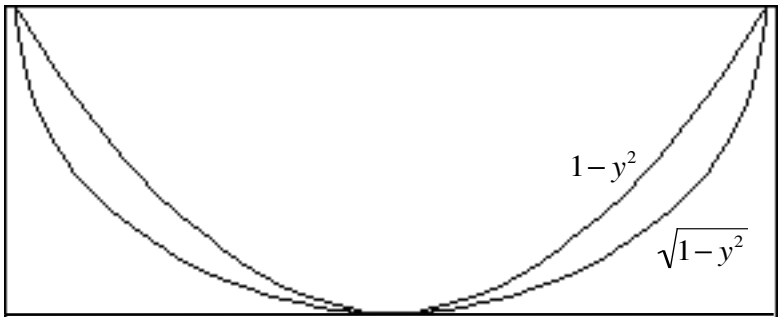
This rotational profile is elliptical. However, as we saw in Chapter 2, real stars should exhibit limb darkening, which will have the effect of reducing the contribution made by large *y* values. This adds another term to the profile of order $1 - y^2$, which is parabolic. The resulting rotational profile is more complex:

Eq. 6.77 $$\phi_{rot}(\Delta\nu) = \frac{3}{3+2\beta}\left[\frac{2}{\pi}\sqrt{1-y^2} + \frac{\beta}{2}\left(1-y^2\right)\right]$$

As before, $y = \dfrac{c\Delta\nu}{\mathrm{v}_e \nu \sin i}$, and β is the first-order limb-darkening coefficient. This is a parabolic-elliptical profile, with the exact form depending on the amount of limb-darkening. Figure 6.14 shows the two components of the rotational profile. The combined profile (Equation 6.77) lies between these two curves.

Rotational broadening is more important in early stars (O, B, and A), which have considerably higher rotational velocities (up to a few hundred km/s) than late stars ($\mathrm{v}_e \approx 2$ km/s for the Sun).

Figure 6.14: The rotational profile components – parabolic ($1-y^2$) and elliptical ($\sqrt{1-y^2}$)



Equivalent Width

The various broadening mechanisms serve to shift line absorption from the transition frequency to other neighboring frequencies. The line strength is thus dependent not only on the depth of the line profile, but also on the width. We need a simple way to quantify this line strength.

We begin by distinguishing between the **line** and the **continuum**: the line opacity $\kappa_{\nu\ell}$ is different from the continuum opacity $\kappa_{\nu c}$, and the line flux F_ν varies across the line while the continuum flux F_c is essentially constant. These fluxes can be used to define the line profile as one of two functions:

Eq. 6.78 $$R_\nu = \frac{F_c - F_\nu}{F_c} = 1 - \frac{F_\nu}{F_c} = 1 - r(\nu)$$

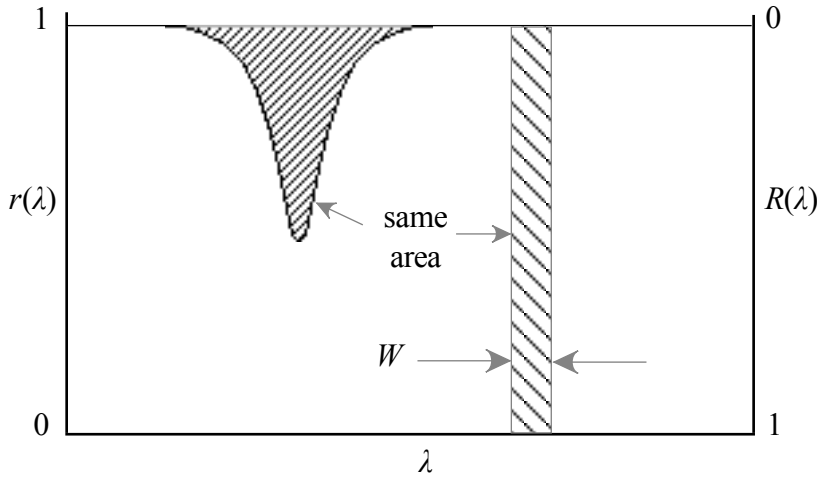
As a measure of line strength we now define the **equivalent width** as follows:

Eq. 6.79 $$W_\nu \equiv \int_0^\infty R_\nu d\nu = \int_0^\infty \left(1 - r(\nu)\right) d\nu \qquad \text{or}$$

Eq. 6.80 $$W_\lambda \equiv \int_0^\infty \left(1 - r(\lambda)\right) d\lambda$$

The equivalent width, illustrated in Figure 6.15, is measured in wavelength units – Ångstroms or milli-Ångstroms. It provides a way to compare line strengths for lines that may have quite different profiles due to different broadening mechanisms.

Figure 6.15: Equivalent width



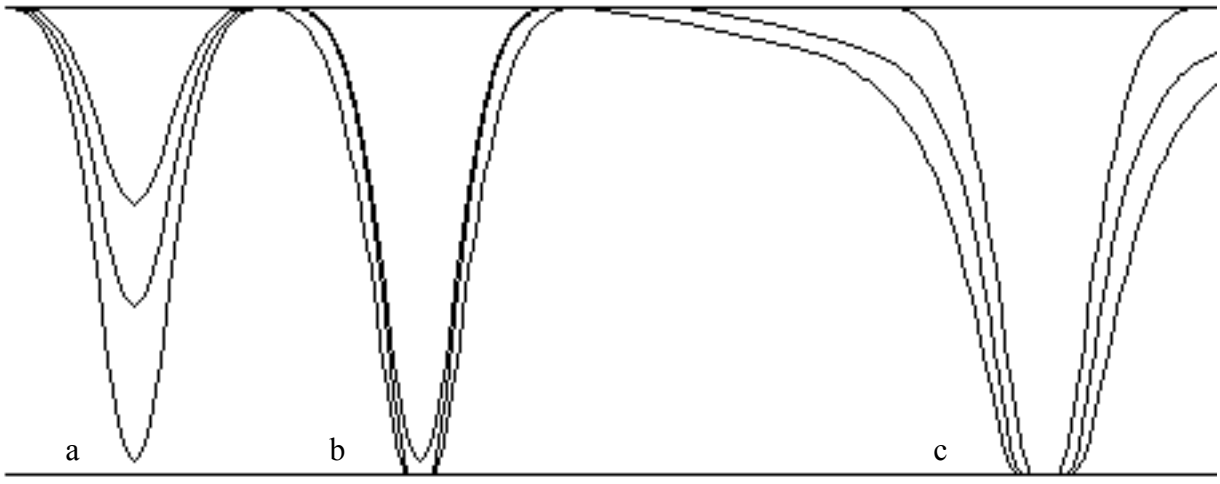
The line strength depends on the absorption coefficient for the line and the **column density** of absorbers, which is the integrated product of the number density and the column height:

Eq. 5.3 $$n_c \approx \int_R^\infty N(r)dr$$

As previously noted, strong lines form high in the atmosphere where temperatures and source functions are lower, making the line appear darker; weak lines are formed deeper in the atmosphere at higher temperatures and source functions.

Examining a weak line, we will normally find a Doppler core. As we increase the column density for this line, more absorption occurs, the Doppler core deepens, and the equivalent width increases in proportion to the column density ($W \propto n_c$), as shown in Figure 6.16a.

Figure 6.16: Growth of a line with increasing column density



As the column density increases further, the core deepens and reaches to the zero flux level. All photons at the transition frequency are then absorbed, and the line appears *black*; at this point the line is said to be **saturated**. Additional absorbers have no more photons to absorb at the

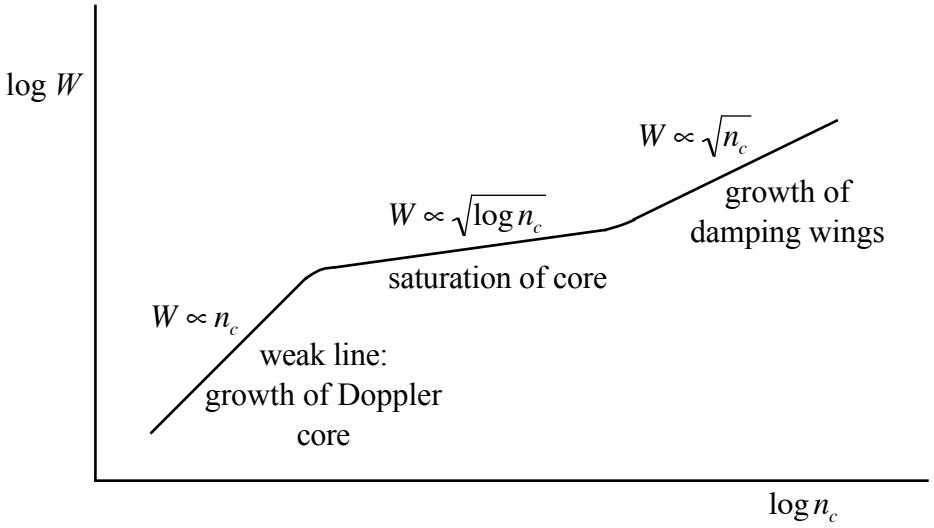
transition frequency, and the equivalent width increases very slowly with column density $\left(W \propto \sqrt{\log n_c}\right)$, as shown in Figure 6.16b.

Eventually the damping wings grow large enough to increase the equivalent width more rapidly $\left(W \propto \sqrt{n_c}\right)$, as shown in Figure 6.16c.

Curve of Growth

A plot of the variation of equivalent width with column density is called the **curve of growth**. Measured equivalent widths can thus be linked to column densities and used to determine abundances, excitation temperatures, ionization degrees, electron pressures, oscillator strengths, etc.

Figure 6.17: The curve of growth



The next chapter discusses the principal sources of opacity that should be considered in analyzing the spectra of stars.

CHAPTER 7: Opacity Sources

There are several mechanisms by which matter may absorb or scatter photons, and these result in different types of opacities with different frequency dependences. Sometimes the difference results from the type of particle involved in the interaction, and sometimes from the manner in which the particle interacts. In general, we may distinguish between **absorption** – in which the energy of the photon is absorbed by the matter – and **scattering** – in which the photon is merely redirected, without significant loss of energy. For each process there are several different mechanisms to understand.

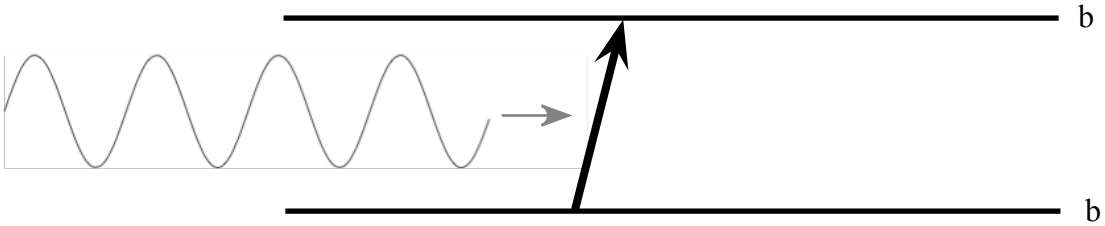
Absorption Transitions

There are three basic types of absorption transitions to consider: **bound-bound** transitions, **bound-free** transitions, and **free-free** transitions.

Bound-Bound Transitions

Bound-bound transitions refer to those in which the electron jumps from one bound state to another. Such transitions between levels of reasonably well-defined energies produce **absorption lines** at distinct frequencies. So far our discussion has focused on this type of transition.

Figure 7.1: A bound-bound transition

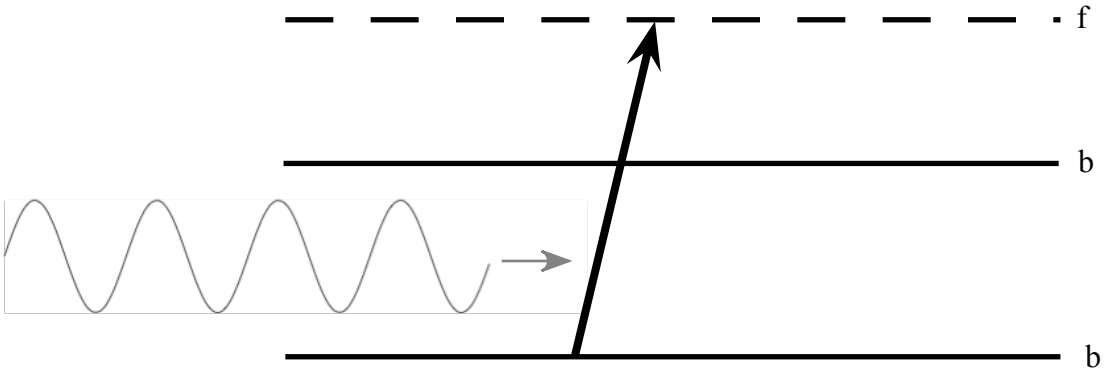


The absorption lines seen in stellar spectra are the result of bound-bound transitions. This mechanism is not an important opacity source at high temperatures (recall the spectra of O stars), but it can become dominant at low temperatures and high densities. The absorption coefficient κ_ν is generally difficult to determine over a range of frequencies because it is the sum of all the individual line opacities.

Bound-Free Transitions

Bound-free transitions involve an electron moving from a bound state within the atom to become a free electron in the gas. This is the process of ionization, or more specifically, **photoionization**, because the transition energy is supplied by the absorbed photon. The energy required is the ionization energy plus whatever kinetic energy the free electron acquires; because this latter amount may be *any* positive energy, bound-free absorption takes place over a wide range of photon energies, producing continuous absorption over a range of frequencies.

Figure 7.2: A bound-free transition

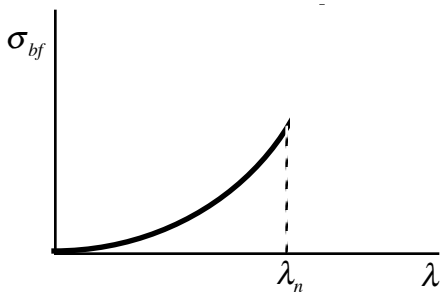


To obtain the frequency dependence of bound-free transitions, we consider a hydrogen-like atom. For such an atom with its electron in level n, all photons with $E \geq \chi_n$ will be capable of producing ionization. This results in a limiting value (λ_n) on the wavelength of a photon that may be absorbed by an atom in level n. The bound-free cross section then takes on the following form:

Eq. 7.1 $$\sigma_{bf} \approx \frac{\lambda^3}{n^5} \text{ for } \lambda \leq \lambda_n$$

A plot of the bound-free cross section vs. wavelength will exhibit sharp **edges** for a given atom.

Figure 7.3: A bound-free absorption edge (at λ_n)



For hydrogen, the limiting wavelengths can be found from equations developed in Chapter 3:

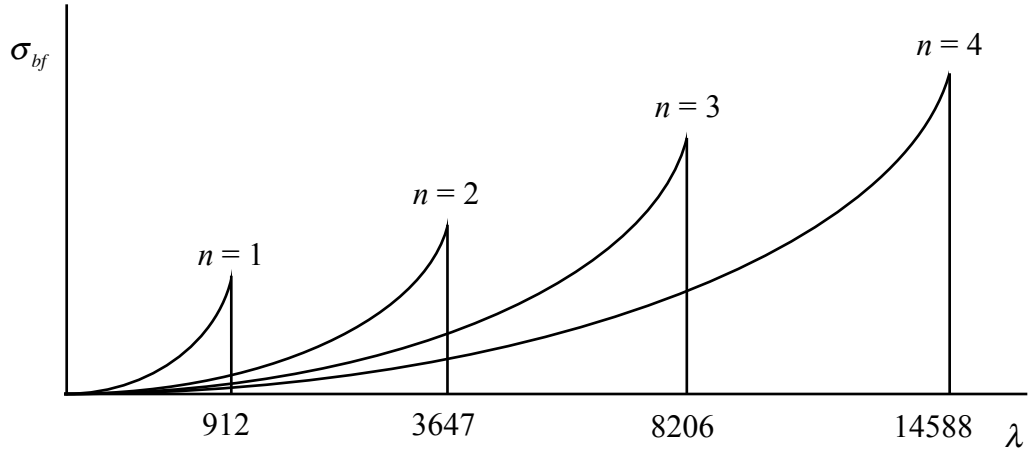
Eq. 7.2 $$\lambda_n = \frac{n^2}{R_H} = 911.7633 n^2 \text{ Å}$$

This yields the following limits for the various series:

- $\lambda_1 = 911.76$ Å (the Lyman series limit)
- $\lambda_2 = 3647.1$ Å (the Balmer series limit)
- $\lambda_3 = 8205.9$ Å (the Paschen series limit)
- $\lambda_4 = 14588$ Å (the Brackett series limit), etc.

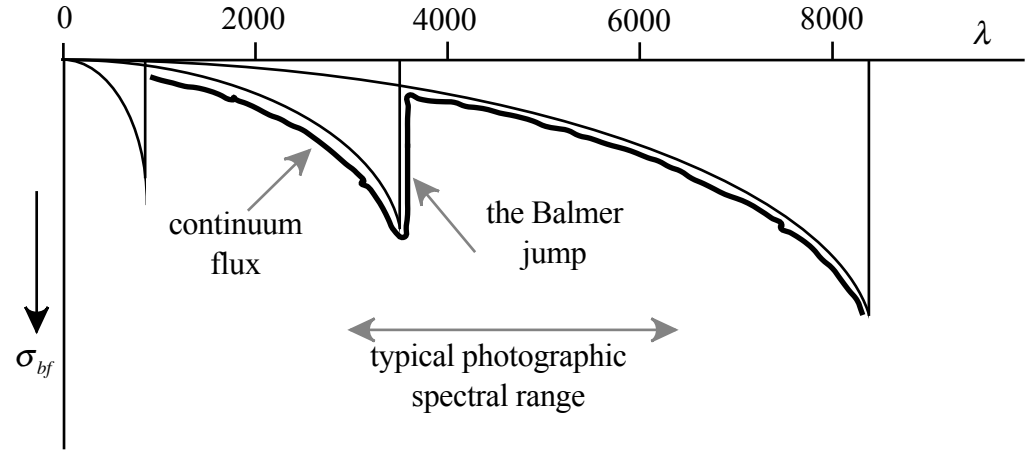
At *visible* wavelengths (4000 to 7000 Å), Paschen absorption should be dominant:

Figure 7.4: Hydrogen bound-free absorption edges



The total bound-free cross section for hydrogen will be a sum over these different series, weighted by the level populations, which are dependent on temperature. At low temperatures, the higher energy levels are less populated, and they play a less important role in the bound-free opacity.

Figure 7.5: The Balmer jump



If hydrogen bound-free absorption is the principal opacity source in a star's atmosphere, then the continuum flux should reflect the wavelength dependence of this source. A sketch of this expected flux is shown in Figure 7.5.

There is a pronounced change in the flux level that should occur just short of 4000 Å, due to the absorption edge associated with the Balmer series limit. This is called the **Balmer jump** or the **Balmer discontinuity**. It can be seen in the spectra of stars that (1) produce significant flux at the ultraviolet wavelengths where the jump occurs, and (2) have a temperature capable of exciting a sufficient number of hydrogen atoms to the second level without completely ionizing them. Thus, the Balmer jump is typically observed in the spectra of B, A, and F stars.

Of course there are other species besides hydrogen that contribute to bound-free absorption, and we will want to include them all – at least all of the significant species. To do this we must combine the various σ_{bf} values into one κ_{bf} term. The problem is as follows:

We have $\sigma_{bf} = \dfrac{\text{cm}^2}{\text{atom } z \text{ in ionization stage } i \text{ \& excitation state } n}$.

We need $\kappa_{bf} = \dfrac{\text{cm}^2}{\text{gram of stellar material}}$.

So we will multiply $\sigma_{bf}(\lambda)\left(\dfrac{\text{cm}^2}{z,i,n}\right)$

by $\left(\dfrac{N_n}{\Sigma N_n}\right)_{i,z}\left(\dfrac{\text{atoms in level } n}{\text{total atoms } z \text{ in } i}\right)$ [Boltzmann equations]

and by $\left(\dfrac{N_i}{\Sigma N_i}\right)_{z}\left(\dfrac{\text{atoms in ionization stage } i}{\text{total atoms } z}\right)$ [Saha equations]

and by $\dfrac{1}{A_z} N_A \left(\dfrac{\text{mole } z}{\text{grams } z}\right)\left(\dfrac{\text{atoms } z}{\text{mole } z}\right)$

and by $X_z\left(\dfrac{\text{grams } z}{\text{gram of *}}\right)$.

Then we will sum over n, i, and z to get κ_{bf}:

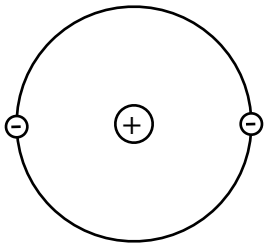
Eq. 7.3 $$\kappa_{bf}(\lambda) = \sum_z \left[X_z \frac{N_A}{A_z} \sum_i \left[\left(\frac{N_i}{\Sigma N_i} \right)_z \sum_n \left[\left(\frac{N_n}{\Sigma N_n} \right)_{i,z} \sigma_{bf}(\lambda, z, i, n) \right] \right] \right]$$

Although hydrogen is the most abundant element in most stars, $\sigma_{bf}(H)$ is dominant only in B, A, and F stars. However, hydrogen does play an important role in the atmospheres of cooler stars through a species called the **negative hydrogen ion *(H^-)***, depicted in Figure 7.6.

As we saw in Chapter 3, the first Bohr orbit can hold two electrons, as long as they have different spins, and a hydrogen atom is just barely capable of accomplishing this feat. The negative hydrogen ion has only one bound state known, meaning that excitation is impossible and no bound-bound transitions may occur. However, photoionization can easily remove the

extra electron from orbit as the ionization energy is only 0.754 eV. This process ($H^- \rightarrow H + e^-$) is designated as $\sigma_{bf}(H^-)$ and all photons with wavelengths less than 16,450Å may participate.

Figure 7.6: Bohr model of the negative hydrogen ion (H^-)



In order for H^- to function as an opacity source, it first must be able to form; this requires neutral hydrogen atoms and a supply of free electrons. But the electrons cannot come from the hydrogen, for the bulk of it must remain neutral (which requires relatively low temperatures). Instead, the electrons must come from species that ionize more readily than hydrogen at low temperatures: the metals.

H^- absorption is thus very weak in hot stars, because its ionization equilibrium lies far to the right. It is weak in very cool stars due to the lack of free electrons caused by insufficient ionization of metals. But in intermediate stars of spectral type G, $\sigma_{bf}(H^-)$ is the dominant source of opacity.

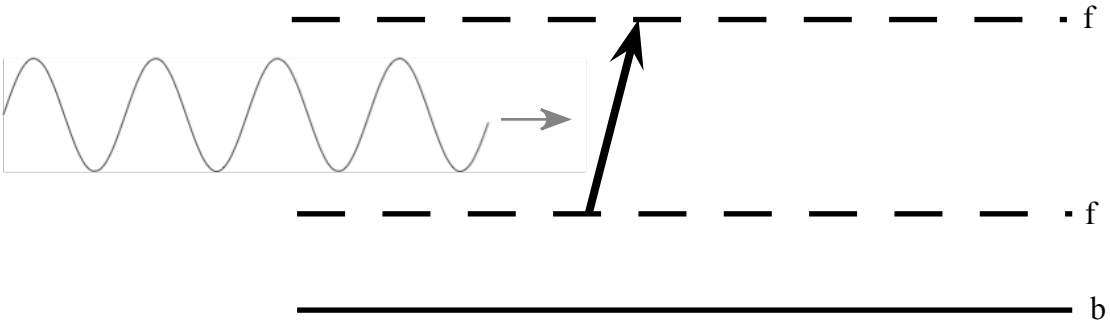
Because $\sigma_{bf}(H^-)$ depends heavily on the presence of free electrons, the cross section is usually given in terms of the electron pressure, as in Figure 7.7:

Figure 7.7: Free-free absorption cross section for H^-

Free-Free Transitions

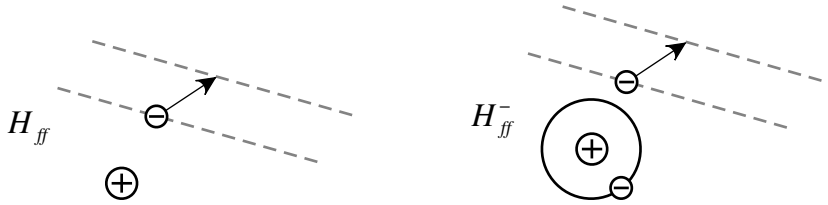
Free-free transitions – also known as **Bremsstrahlung** – occur when a free electron in the vicinity of an ion absorbs a photon, changing the electron's energy relative to the ion, as shown in Figure 7.8. No bound states are involved. The cross section for free-free transitions depends on the effective atomic number of the ion and on the speed of the electron (that is, on the temperature of the gas).

Figure 7.8: A free-free transition



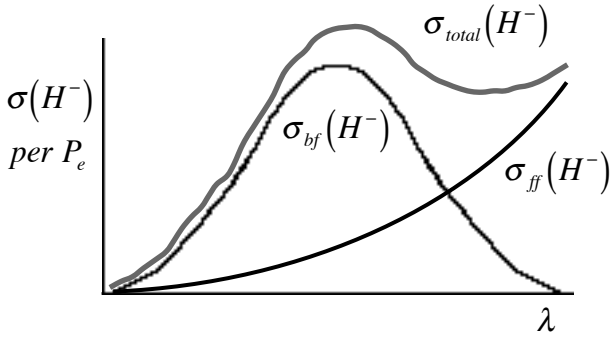
For interactions with hydrogen nuclei (protons), we find the free-free cross section varies with the cube of the wavelength: $\sigma_{ff}(H) \approx \lambda^3$. The electron can also interact with a hydrogen *atom*, producing what is known as $\sigma_{ff}(H^-)$ (even though there is no H^- actually formed). Figure 7.9 illustrates the difference.

Figure 7.9: Hydrogen free-free transitions



Normally H^-_{bf} and H^-_{ff} are considered together, as shown in Figure 7.10. H^-_{ff} absorption becomes dominant at infrared wavelengths.

Figure 7.10: Total absorption cross section for H^-



Scattering

Scattering refers to the deflection of photons by matter. While the photon's *direction* is normally altered by scattering, its *energy* is usually unchanged. The scattering particle may be a free electron, an atom, an ion, a molecule, etc. The type of scattering depends on the energy of the photon and the resonant wavelength of the scattering particle.

Photons can be scattered by charged particles, primarily electrons. We may envision the oscillating electric field of the wave inducing similar oscillations in the charged particle, and the resulting acceleration causes the charge to radiate. As this radiation occurs over random directions, the incoming wave is effectively scattered. This classical explanation of scattering by a charge is known as **Thomson scattering**, or **electron scattering**, as electrons are the principal particles involved.

Thomson Scattering

The cross section for Thomson scattering σ_T can be derived classically; it is given here in terms of the classical electron radius r_e, which is obtained by equating the electron's electrostatic potential energy to its rest energy:

Eq. 7.4 $$\frac{e^2}{r} = mc^2 \quad \Rightarrow \quad r_e = \frac{e^2}{m_e c^2} = 2.829 \times 10^{-13} \text{ cm}$$

A classical analysis then produces the **Thomson cross section**:

Eq. 7.5 $$\sigma_T = {}^8\!/_3 \, \pi r_e^{\,2} = 6.706 \times 10^{-25} \text{ cm}^2\text{/electron} \approx {}^2\!/_3 \text{ barn} \qquad \text{where 1 barn} = 10^{-24} \text{ cm}^2$$

Thomson scattering will be important in stellar interiors and in the atmospheres of hot stars, where ionization of hydrogen produces copious amounts of free electrons. (One might wonder about considering similar scattering off the positively charged hydrogen nuclei (protons) as well; however, because $\sigma \approx {}^1\!/_{m^2}$ and ${}^1\!/_{m_p^2} \approx 3 \times 10^{-7}$ $({}^1\!/_{m_e^2})$, the effect is considerably smaller, and contributions from nuclei may be safely ignored.)

This explanation of Thomson scattering is valid as long as the energy of the photon is small compared to the rest energy of the electron:

Eq. 7.6 $$h\nu << m_e c^2 \quad \Rightarrow \quad {}^{hc}\!/_{\lambda} << m_e c^2 \quad \Rightarrow \quad \lambda >> {}^h\!/_{m_e c} \equiv \lambda_c \approx 0.0243 \text{ Å}$$

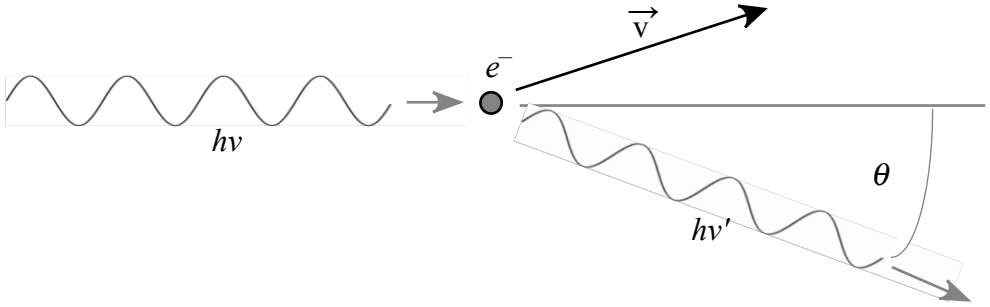
The limiting value on the wavelength is the **Compton wavelength**, λ_c. This condition will normally be met by free electrons at temperatures less than about 10^9 K (set $kT \approx m_e c^2$), which covers most stellar interiors.

A photon of higher energy – such that $\lambda = \lambda_c$ or less (X-rays and gamma rays) – may be able to *accelerate* the electron (rather than just bouncing off it), transferring some of its energy and momentum to the electron, and thus diminishing the energy of the scattered photon. This requires a quantum treatment of the photon as a *particle*, rather than a wave, resulting in a process called **Compton scattering**.

Compton Scattering

Compton scattering occurs when a high-energy photon scatters off a stationary electron, losing some energy in the process to the electron. The scattering angle is shown in Figure 7.11.

Figure 7.11: Compton scattering



The change in photon energy depends on the scattering angle θ :

Eq. 7.7 $$h\nu' = \frac{h\nu}{1+\frac{h\nu}{m_e c^2}(1-\cos\theta)} \Rightarrow \frac{hc}{\lambda'} = \frac{\frac{hc}{\lambda'}}{1+\frac{h}{m_e c\lambda}(1-\cos\theta)}$$

This provides the wavelength of the scattered photon:

Eq. 7.8 $$\lambda' = \lambda\left(1+\frac{h}{m_e c\lambda}(1-\cos\theta)\right) = \lambda + \frac{h}{m_e c}(1-\cos\theta)$$

And the *change* in wavelength is as follows:

Eq. 7.9 $$\Delta\lambda = \lambda' - \lambda = \lambda_c(1-\cos\theta) = 2\,\lambda_c \sin^2(\theta/2)$$

Again, $\lambda_c \equiv h/m_e c = 0.0243$ Å is the Compton wavelength for the electron. Because this value is so small, the change in the energy of a scattered photon will be significant only for high energy photons, with $\lambda = \lambda_c$ or less. Low energy photons are simply scattered over 4π steradians with essentially no loss of energy (Thomson scattering).

Rayleigh Scattering

Photons may also be scattered off bound electrons in atoms and molecules. For these interactions, we must require that the photon energy be less than the energy difference between any adjacent bound states ($h\nu < \Delta E_{nm}$) in order that no transition occurs. Given this restriction, the cross section for Rayleigh scattering can be written in terms of the Thomson cross section:

Eq. 7.10 $$\sigma_R = \sigma_T \frac{1}{\left[(\nu_o/\nu)^2 - 1\right]^2} \qquad \text{where } h\nu_o \approx \Delta E_{nm}$$

Now for $h\nu << \Delta E_{nm}$, $\nu << \nu_o$, and $\nu/\nu_o << 1$. This yields the normal approximation for Rayleigh scattering:

Eq. 7.11 $$\sigma_R = \sigma_T \frac{(\nu/\nu_o)^4}{\left[1-(\nu/\nu_o)^2\right]^2} \approx \sigma_T \left(\frac{\nu}{\nu_o}\right)^4 = \sigma_T \left(\frac{\lambda_o}{\lambda}\right)^4$$

The result is that there is a strong wavelength dependence for Rayleigh scattering, such that short wavelength photons are scattered much more readily. This of course is responsible for the blue skies we observe in our atmosphere, where visible photons from the Sun scatter off the electrons in nitrogen and oxygen molecules in the air.

As a source of stellar opacity, Rayleigh scattering can become important at the lower temperatures found in the atmospheres of cool stars. There, the most likely targets are neutral hydrogen atoms, giving a cross section as follows:

Eq. 7.12 $$\sigma_R = \sigma_T \left(\frac{\lambda_L}{\lambda} \right)^4$$

Here, λ_L = 1026 Å, which is a weighted average of the Lyman wavelengths – used because most of the hydrogen atoms should be in the ground state at these temperatures. Again, the limits on Rayleigh scattering require that $\lambda > \lambda_L$, which means that the Rayleigh cross section will be *smaller* than the Thomson cross section. However, at these low temperatures, free electrons are not particularly plentiful, and Rayleigh scattering may be dominant over Thomson scattering.

Novotny (1973, pp 136-150) provides figures that combine all these sources to give opacities for stars of different temperatures.

In the next chapter we will try to deduce the structure that results when matter and radiation interact within the stellar atmosphere.