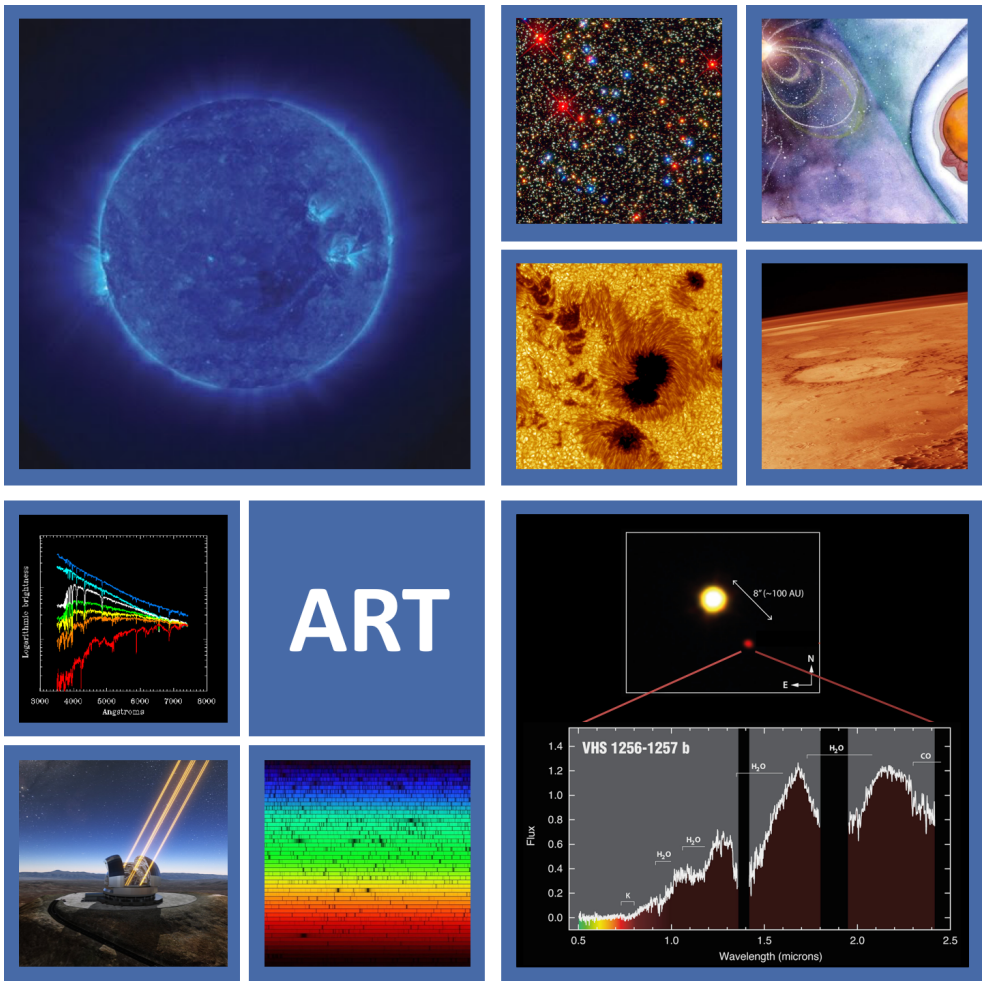


Atmospheres & Radiative Transfer

Prof. dr. Alex DE KOTER

Astronomical Institute *Anton Pannekoek*
University of Amsterdam
P.O. Box 94249, 1090 GE Amsterdam,
The Netherlands

Instituut voor Sterrenkunde,
KU Leuven
Celestijnenlaan 200D, 3001, Leuven,
Belgium

Version 1.103.00

Course objectives

To learn the basic background of radiative transfer processes in astrophysical media, with special emphasis on the analysis of stellar atmospheres. More specifically:

- ☐ The student can give a description of the Morgan-Keenan (MK) classification of stellar spectra, explain the underlying physical ideas, and classify an optical stellar spectrum as an early-, mid- or late-type spectrum.
- ☐ The student is aware that a small fraction of stars cannot be classified according to MK, can explain the nature of these stars and the reasons why they require a separate classification scheme
- ☐ The student can describe the Johnson photometric system and can give an expose on the pros and cons of photometric versus spectroscopic observations in a given astrophysical case.
- ☐ The student knows the definitions of the main radiation field quantities, understands their physical meaning, and can compute their values in given astrophysical cases.
- ☐ The student can prove that specific intensity is distance invariant and that flux decreases with the square root of the distance.
- ☐ The student can give a heuristic derivation of the equation of transfer.
- ☐ The student can solve the equation of transfer for source functions that are linear in optical depth.
- ☐ The student can analyse the optically thick and optically thin limit of radiative transfer problems.
- ☐ The student can explain the basic physics of line formation in a static atmosphere.
- ☐ The student understands the different assumptions underlying TE, LTE and NLTE and can give an expose on which of these equilibrium states is most representative in a given astrophysical case.
- ☐ The student understands the basics of the methodology to derive the excitation and ionization state of a gas in LTE.
- ☐ The student can give an expose on different temperature definitions
- ☐ The student understands the nature of, and can give an expose on, opacities resulting from both discrete and continuous processes, and is able to identify the main sources of extinction in stellar and planetary atmospheres.
- ☐ The student understands the physical meaning of the conservation laws and can derive the density structure and total mass of a hydrostatic isothermal atmosphere.

- ☐ The student can explain how the basic global properties (mass, luminosity and radius) and basic atmospheric properties (effective temperature, surface gravity, abundances, micro-turbulent velocity, and projected surface rotation velocity) of a star can be constrained and what kind of observations are required to do this.
- ☐ The student can synthesize his/her basic knowledge of radiative transfer to build a grey planar LTE atmosphere in hydrostatic and radiative equilibrium.
- ☐ The student understands the concept of constructing LTE and NLTE model atmospheres and can explain why scatterings are the main problem in building such models.
- ☐ The student can give an expose on spectral line-broadening mechanisms and can explain the physical processes responsible for these line-broadening processes.
- ☐ The student can give an expose on the curve of growth method.
- ☐ The student can give an overview of the structure of the solar atmosphere.
- ☐ The student can explain the basic physics of line formation in outflowing atmospheres and the formation of P Cygni, flat-topped and parabolic line profiles.
- ☐ Using the lecture notes, the student is able to derive the relations needed to constrain the mass-loss rate from the Halpha spectral line and radio continuum flux.
- ☐ The student can describe the impact of inter-stellar gas and dust on the spectra of stars, and can give an expose on the potential problems such contaminations may cause when interpreting stellar and planetary spectra.
- ☐ The student can apply the acquired skills in radiative transfer to problems other than those related to atmospheres.

Physical Constants

Name	Symbol	CGS Value	CGS units
Speed of light in a vacuum	c	2.998×10^{10}	$\mathrm{cm\,s^{-1}}$
Planck constant	h	6.626×10^{-27}	erg s
	$\hbar$	1.055×10^{-27}	erg s
Gravitational constant	G	6.673×10^{-8}	$\mathrm{cm^3\,g^{-1}\,s^{-2}}$
Atomic mass unit	$\mathrm{m_{amu}}$	1.661×10^{-24}	g
Mass of hydrogen	$\mathrm{m_H}$	1.673×10^{-24}	g
Mass of proton	$\mathrm{m_p}$	1.673×10^{-24}	g
Mass of neutron	$\mathrm{m_n}$	1.675×10^{-24}	g
Mass of electron	$\mathrm{m_e}$	9.109×10^{-28}	g
Avagadro's number	N_A	6.022×10^{23}	
Gas constant	$\mathcal{R}$	8.314×10^{7}	$\mathrm{erg\,K^{-1}\,mol^{-1}}$
Boltzmann constant	k	1.381×10^{-16}	$\mathrm{erg\,K^{-1}}$
Electron volt	eV	1.602×10^{-12}	erg
Radiation density constant	a	7.565×10^{-15}	$\mathrm{erg\,cm^{-3}\,K^{-4}}$
Stefan-Boltzmann constant	σ	5.671×10^{-5}	$\mathrm{erg\,cm^{-2}\,K^{-4}}$
Fine structure constant	α	7.297×10^{-3}	
Rydberg constant	R_∞	2.180×10^{11}	erg

Astronomical Constants

Name	Symbol	CGS Value	CGS units
Astronomical unit	AU	1.496×10^{13}	cm
Parsec	pc	3.086×10^{18}	cm
Light year	ly	9.463×10^{17}	cm
Solar mass	$M_\odot$	1.989×10^{33}	g
Solar luminosity	$L_\odot$	3.828×10^{33}	$\mathrm{erg\,s^{-1}}$
Solar radius	$R_\odot$	6.957×10^{10}	cm
Solar effective temperature	$T_{\mathrm{eff},\odot}$	5 772	K
Thomson scattering coefficient	σ_T	6.652×10^{-25}	$\mathrm{cm^2}$

Conversion Factors

Name	Symbol	Value	CGS units
Year	yr	3.156×10^{7}	s
Arcsec	”	4.848×10^{-6}	radians
Solar mass per year	$M_\odot \mathrm{yr}^{-1}$	6.303×10^{25}	$\mathrm{g\,s^{-1}}$

Contents

1

Introduction

In the vast majority of cases radiation is the only information we have from distant astronomical objects. Almost all knowledge we have gathered about stars, planets and moons, the interstellar medium, galaxies, et cetera, is derived from analysis of this radiation. It is therefore of considerable importance to study the physics of the creation and transport of radiation in these objects, and to develop reliable methods that allow to decode the information about such systems that is contained in their spectra.

If we construct a model of the problem of interest and predict the corresponding spectral energy distribution, we can make quantitative statements about the physical state of the surface layers of the object that is observed, if we compare our predictions with the measurements. In this way we may learn about properties such as temperature and density structure, the kind of gases and solid particles present, magnetic field strengths and topologies, velocity fields, force fields (particularly that due to gravity), mass, mass loss and mass gain, luminosity of the source as a whole, geometrical properties (e.g. radius in the case of an opaque spherical object), and distance.

1.1 Radiation as probe and as constituent

In the astrophysical context, radiation has a conspicuously dualistic character. On the one hand it carries the information that we use to derive the physical state of a medium, i.e. it serves as a probe of this medium. On the other hand it is, in many cases, an important constituent of this medium, i.e. it in fact determines the structure of the medium itself.

A well-known example of such an astrophysical medium is the stellar atmosphere. In the outermost layers of a star photons escape to the surrounding space, taking away energy. This energy is often transported from the near-surface layers to the edge by photons, and photons therefore in part determine the temperature, and – because of radiation pressure – density of the atmospheric layers. If radiation pressure is intense, photons may drive a stellar wind through which matter is lost to space. Coded in the spectral lines and the continuum energy

distribution, that both originate in the stellar atmosphere, is information on the local conditions. This illustrates the dualistic nature of radiation.

In the past decades impressive progress has been made in the quantitative analysis of atmospheres. Several reasons for this may be identified: *i)* the unprecedented increase in the quality of ground- and space-based observations; *ii)* the development of extremely fast and efficient numerical methods; and *iii)* an ever increasing computer power. The theory of stellar atmospheres has reached such a high level of sophistication that it can be used as a methodological handbook for the study of other astrophysical systems where radiation has the above mentioned dualistic nature. Examples of such systems are planetary atmospheres, the circumstellar medium, accretion disks, H II regions, planetary atmospheres and rings, and the interstellar and intergalactic medium.

In these lectures we focus on the physics of the creation and transport of radiation, and on methods that help us decode spectral information. We primarily discuss stellar and planetary atmospheres, however several chapters are devoted to other astrophysical objects.

1.2 About these lecture notes

In chapter 2 we discuss the spectra of stars, brown dwarfs, and planets and their classification. This classification is done primarily on the basis of spectral line properties, but may also be done using broadband color indices. The calibration of the spectral types, luminosity classes and color indices in terms of effective temperature, gravity and chemical composition is done using models of stellar atmospheres. The first 8 chapters are devoted to discussing the basic physics of radiative transfer and of extinction and emission. This knowledge is applied in chapters 9 through 13 (and in the optional chapters 14 and 15) to construct model atmospheres. In Chapter 16 we discuss our star, Sun. The final chapters discuss other astrophysical media.

In chapter 3 we introduce the macroscopic quantities that characterise the radiation field; in chapter 4 the equation that describes the transfer of radiation is derived. This 4th chapter also discusses simple analytical solutions of the radiation transfer equation. Numerical solutions are treated in (the optional) chapter 5. The coupling between the radiation field and the material medium is discussed in chapter 6. We will invest time to study types of equilibrium between radiation and matter. In thermodynamic equilibrium (TE) the state of the gas is fully described by the equations of Boltzmann and Saha. The Saha equilibrium specifies the ionisation as a function of temperature and density of the medium. This provides a relatively simple, but already quite successful way to classify a stellar or planetary spectrum (see § 6.5). In TE the radiation field is described by the Planck function. This is the simplest description of the emerging radiation field of a star or planet, and could, at least in principle, be used to calibrate color indices (see § 6.6). However, a real stellar spectrum is more complex and shows appreciable deviations from that of a Planck curve. A meaningful calibration therefore requires more sophisticated models.

The transfer of radiation is dependent on the sources of extinction χ_ν and of emission η_ν in the medium. In chapters 7 and 8 we make the link between these macroscopic quantities and the microscopic processes that cause extinction and emission.

Starting with chapter 9 we focus on the modelling of atmospheres. We start out with a discussion of the conservation of mass, momentum, and energy. In essence, the problem of stellar and planetary atmospheres is to solve the equation of radiation transfer for a suitable equilibrium between radiation and matter, subject to the conservation constraints given above.

In chapter 10 we discuss the grey atmosphere. This relatively simple model provides important insight in the role of the energy equation. A widely used model to describe the stellar atmosphere assumes that the gas is in local thermodynamic equilibrium (LTE) and that hydrostatic and radiative equilibrium hold. This type of model is treated in chapter 11. In chapter 12, we focus in on planetary atmospheres. Next, the processes that determine the shape and strength of spectral lines are reviewed. These need to be included in our models if the aim is to construct a representation of the atmosphere that may be compared to observations. Chapter 14 focusses on the process of scattering, arguably the most fundamental *physical* complication inherent to solving the model atmosphere problem. In some cases, the assumptions that make up LTE break down in the stellar atmosphere. In that case non-LTE models are needed to describe the medium. Aspects of non-LTE physics & models are discussed in chapter 15.

The sun is such a special star that we devote a separate chapter to it. In the final chapters we 'move away from the stellar surface'. We successively discuss stellar winds, nebulae, and the interstellar and intergalactic medium.

Exercise 1.1

a) Give astrophysical information carriers other than electromagnetic radiation.

b) Think of types of observations and areas of astronomical research where knowledge of the creation and transport of radiation is *not* important.

1.3 Literature

No books fully cover what is presented in these lecture notes. However, we follow parts of:

- Mihalas: *Stellar Atmospheres*, 2nd edition. W.H. Freeman and Company, New York, 1978. Standard work, including a detailed description of the basis physics of stellar atmospheres.

- Hubeny & Mihalas: *The Theory of Stellar Atmospheres*, 1st edition. Princeton University Press, 2014. Essentially the 3rd edition of the book of Mihalas, completely reworked by Hubeny to include modern techniques. Standard work.
- Rybicki and Lightman: *Radiative Processes in Astrophysics*. Excellent standard book. Treats parts of the basic physics in more detail than is done in these lecture notes.
- Pradhan & Nahar: *Atomic Astrophysics and Spectroscopy*, 1st edition. Cambridge University Press, 2011. Excellent overview of atomic structure and processes relevant for radiative transfer.
- Gray: *The Observation and Analysis of Stellar Photospheres*, 1st edition. Cambridge University Press, 1992, 2005, 2008. Excellent book. More basic than these lectures, with an emphasis on instrumentation and observing techniques in optical spectroscopy.
- Gray & Corbally: *Stellar Spectral Classification*. Very useful book. Discusses the spectral classification method, but unfortunately does not provide an overview of spectral calibration.
- Rutten: *Opwekking en Transport van Straling*. Second year lecture notes of the University of Utrecht (in Dutch). Very clear and conceptual discussions. Though at a lower level, it is used as the main guide to write several chapters of these lecture notes.
- Rutten: *Radiative Transfer in Stellar Atmospheres*. Doctoral lectures at the University of Utrecht. Almost encyclopedic approach, though less conceptual than his second year lectures. Available through internet. Comparable level.
- Hubeny: In: *Stellar Atmospheres: Theory and Observations*. Excellent discussion of modern techniques in radiation transfer.
- Seager: *Exoplanet Atmospheres*. Clear and concise textbook on exoplanet atmospheres, with a brief introduction of basic concepts of radiative transfer.
- Osterbrock & Ferland: *Astrophysics of Gaseous Nebulae and Active Galactic Nuclei*. University Science Books, California, 2006, Second Edition. Standard textbook on the physics of nebulae and AGNs. Beautifully written.
- Aller: *Physics of Thermal Gaseous Nebulae*. Standard work on the physics of nebulae with many useful tables. Well written though printed in an horrific font.
- Lamers & Cassinelli: *Introduction to Stellar Winds*. Standard work on the theory of stellar winds. Clear, very didactic approach.
- Aller: *The Atmospheres of the Sun and Stars*. Somewhat outdated, though very readable book on the physics of stellar spectra and stellar atmospheres with a good introduction on atomic and molecular line formation.
- Hearnshaw: *The Measurement of Starlight*. Very readable book on the history of stellar spectroscopy.

2

The spectra of stars, brown dwarfs, and planets

The history of astronomical spectroscopy reaches back for two centuries. In this introductory chapter we first give a short (and therefore incomplete) summary of this history. We then discuss the spectra of stars and planets and their classification.

2.1 Spectral classification

In 1802 William Hyde Wollaston (1766–1828), an English chemist and physicist, found that the spectrum of the sun contained dark stripes (spectral lines), which he interpreted as natural separations between colors. From 1814 on the German maker of lenses Joseph von Fraunhofer (1787–1826) studied these lines in detail, which he labeled with the letters A to K (a notation which in part is still in use today). He hoped to be able to use these dark stripes as markers of wavelengths for the development of optical instruments. Fraunhofer also observed the spectra of stars in the night sky, and noticed large differences in color and stripe pattern. In the yellow and orange part of the spectrum of Sirius and Castor he could not detect lines at all, however, in the green-blue part he found three strong lines. The spectra of Pollux and Capella were very similar to that of Venus (i.e. to that of the sun). The spectrum of Betelgeuse did not look like that of the sun at all, though Fraunhofer thought he saw the D-line.

Figure 2.1: *Joseph von Fraunhofer (1787–1826).*

Later, others described the spectral differences between stars as recognized by Fraunhofer in more detail. These differences were used to classify stars. The first systematic classification is from father Angelo Secchi (1818–1878), who founded a Jesuit observatory in the city of

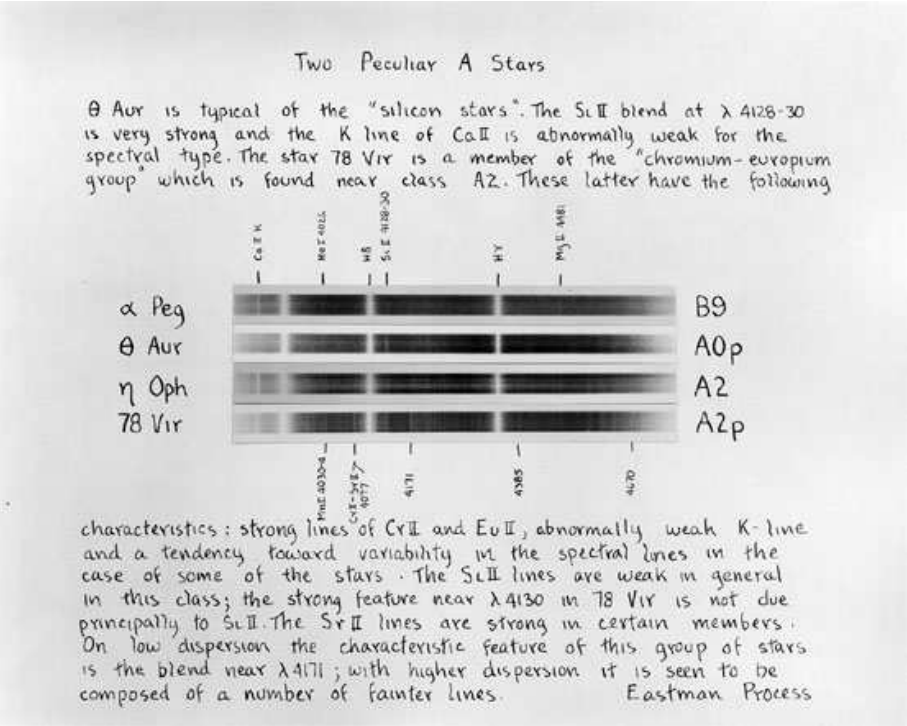


Figure 2.2: *Discussion of the peculiar A-type stars by Morgan, Keenan & Kellman in their 'Atlas of stellar spectra, with an outline of spectral classification' (1943), which is at the basis of the MK classification scheme.*

Rome in the mid of the nineteenth century. In 1867 he published a catalogue of 314 stars subdivided in four types. In 1890 at Harvard Observatory a large program to classify stars was initiated by Edward Charles Pickering (1846–1919), leading to the Henri Draper Memorial Catalogue, which contained the spectra of 10 351 stars. These spectra were recorded on photographic objective-prism plates, and were classified by means of visual inspection. The estimated ratio between the strengths of spectral lines determined their spectral class.

The spectral classes were denoted by the capital letters A through P (later some of these letters disappeared as labels of spectral type). Besides this large scale classification project more detailed studies were done, from which emerged that some stars needed an extra label to properly describe their spectra. For this, small letters were used. Some examples of this are: e, in case emission lines are present, especially those of the Balmer series[1]; c, in case the spectral lines were unusually sharp; f, in case of emission in the He II λ4686 and N III λ4640 lines. The latter may occur in the spectra of O-type stars.

The great importance of a precise classification became clear with, among others, the discov-

[1]Hydrogen lines from main quantum level n (> 2) to 2 are referred to as Balmer lines.

ery by Ejnar Hertzsprung (1873–1967)[2] in 1905 that the proper motion of early-type stars that carried the label c was much smaller than that of normal stars of the same spectral class, having comparable apparent magnitudes. Hertzsprung concluded that c-stars of equal apparent magnitude were much more distant than normal stars, and therefore intrinsically much brighter. We now know that these c-stars are giants and supergiants. The sharpness of their spectral lines is linked to a much lower density of their atmospheres, which results in less 'pressure broadening' of these lines (see chapter 13).

On an even larger scale, the spectra of almost a quarter of a million stars were classified at Harvard and published from 1918 to 1924. This work, by Annie Jump Cannon (1863–1941) and assistants, culminated in a new version of the Henri Draper (HD) Catalogue, in which 225 300 stars were included. Classification of a further 46 850 stare were published in the Henry Draper Extension (HDE) catalogue. In this monumental catalogue the spectral classes were indicated by capital letters, followed by a number (0 to 9) for a possible finer subdivision. The original series A through P was thinned out, and ended up in the non-alphabetic order we use today[3]:

O B A F G K M

← earlier types later types →

This enormous industry was purely morphological. The idea was that the classification reflected the evolution of stars, hence the terms 'early-' and 'late'-type stars. Even after Hertzsprung in 1908 and, independently, Henry Norris Russell (1877–1957) in 1913 presented their diagram of absolute magnitude versus spectral type, the true meaning of the spectral types remained unclear. Only after Anton Pannekoek realized the implications of pressure broadening; Meghnad Saha derived the equations for ionization equilibrium, and Alfred Fowler and Edward Arthur Milne coupled the colors of stars to differences in ionization, things became clear.

The spectral sequence is one of stellar surface temperature. As we will see later in these lectures it is predominantly temperature that determines the ionization and excitation properties of the outer layers of a star, where the spectral lines are formed. The degree of ionization of an element depends also on the density of the medium. This forms the basis of the luminosity classification of stars, first proposed by William Sydney Adams (1876–1956) in 1914, and later refined by many others, and which, after calibration, can be used for 'spectroscopic parallax' determination. In its present form the luminosity-classification divides stars in types

[2]Born in Denmark, Hertzsprung worked at Leiden Observatory from 1919 to 1944, the last nine years as director.

[3]Mnemonics to remember the spectral order are, for the ladies: O Be A Fine Guy, Kiss Me; and for the gentlemen: Only Boys Accepting Feminism Get Kissed Meaningfully.

Table 2.1: *An overview of the Morgan and Keenan (MK) spectral classification scheme, and of additions made to this system. For extremely cool objects, such as very cool stars, brown dwarfs, and planets with gaseous atmospheres the types L, T, and Y have been added. For giants with peculiar abundance patterns the types S, R, N, and C are used.*

MK spectral class	Class characteristics
O	Strong UV continuum with He II absorption. Lines from: C III, N III, O III, Si IV
B	No He II lines. He I absorption, with maximum at B2; H lines stronger at later subtypes. Lines from: C II, O II, Si III
A	H lines reach maximum at A0, weaken at later subtypes. Weak Ca II appearing. Mg II and Si II strong.
F	Ca II stronger; H lines weaken; lines of singly ionised metals prominent; those of neutral metals appear
G	Ca II very strong; strong neutral and weaker ionic metal lines, e.g. of Fe; H lines weaker
K	Blue continuum weak. Strong neutral metallic lines, e.g. Ca I; H lines very weak; CH and CN bands developing
M	Red continuum; very strong Ca I; TiO bands developing strongly.
L	Infrared continuum. Weakening VO & TiO bands; CO, H_2O, CaH, FeH, CrH absorption bands. Strong lines of easily ionised atoms Li, Na, K, Rb, Cs (i.e. alkali metals)
T	Strong infrared continuum. Appearance of strong CH_4, H_2 and NH_3 bands in near-IR
Y	Strong mid-infrared continuum; CH_4, H_2O, and NH_3 bands.

Additional spectral classes	Class characteristics
S	Strong bands of ZrO and YO, LaO present, TiO weak compared to M stars, strong CN bands though not as strong as in C stars
R (or C)	Strong bands of CN and CO instead of TiO in class M
N (or C)	Swan bands of C_2, Na I (D), Ca I λ4227, for the rest similar to R

Additional qualifiers sometimes included with MK types	
e	emission components in hydrogen lines when not expected, e.g. Be
f	N III λ4634−40−42 emission, He II λ4686 emission in certain O type stars (for an even more detailed sub-classification of Of stars, see Sota et al. 2011, ApJS 193, 24)
p	peculiar spectrum, e.g. anomalous line strengths or nebular (diffuse) lines
n	broad lines (nn – even more broadened lines; nnn - extremely broad lines)
s	sharp lines
c	especially sharp lines, characteristic of supergiants
k	interstellar lines present
m	conspicuously strong metal lines (usuall applied to A stars)
v or var	variable lines

MK luminosity class		
I		supergiants
	Ia^+ or Ia-0	extreme supergiants
	Ia	luminous supergiants
	Iab	normal supergiants
	Ib	underluminous supergiants
II		bright giants
III		giants
IV		subgiants
V		dwarfs (or main sequence stars)
VI	or sd (prefix)	subdwarfs
VII	or D (prefix)	white dwarfs

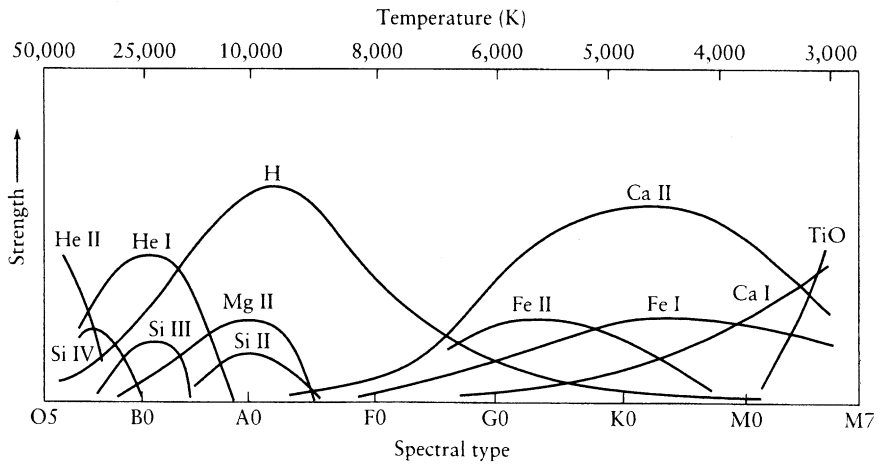


Figure 2.3: *Strength of key features in the spectra of stars as a function of spectral type.*

I, II, III, IV, and V (see figure 2.5). A more refined devision of class I (supergiant) stars can be made using the letters a^+, a, ab, and b (class Ia^+ being the most luminous).

Todays standard for the classification of stellar spectra is that defined by William Morgan, Philip Keenan, and Edith Kellman in the forties, and later revised by Morgan and Keenan (MK), and which is based on spectra of standard stars. The earliest spectral type in this system is O3 (O0..O2 are not used); the latest type is M8.[4] Of all stars brighter than magnitude 8, for which the Henri Draper Catalogue is still complete, 99.95% can be classified in this way. An overview of the MK classification scheme is given in table 2.1. Roughly the classification works as follows (see also Fig. 2.3)

O- and B-type stars are principally classified on the basis of lines of hydrogen and of neutral and singly ionized helium. In addition lines of several ionization stages of heavier elements, such as C, N, O, and Si, are used. For a detailed overview of the O and B star classification see Liu et al. 2019 and Table B.1. *A-type stars* are difficult to classify, as in this group of stars the Balmer series lines reach their maximum strength, and most lines of other elements are relatively weak. One uses the K-line of Ca II, which rapidly increases in strength from B8 to F0, as the prime criterion. *F0 to early M-type stars* show the following changes: In general the F-type stars show the absorption lines of heavy elements. These lines become stronger toward later spectral type. From type F toward later types the Balmer series lines become weaker, and are no longer conspicuous in K- and M-type spectra. Classification of F- to K-type spectra is based on Balmer lines, the lines of Fe I, Ca I, Ca II, and the G-band at ca. 4300 Å (a superposition of lines of the CH molecule and of a number of heavy elements, which Fraunhofer referred to with the letter G). *M-type stars* are characterized by the presence of molecular TiO bands, which rapidly increase in strength toward later types.

[4] More recently also the types O2 and M8.5, 9, and 9.5 have been introduced.

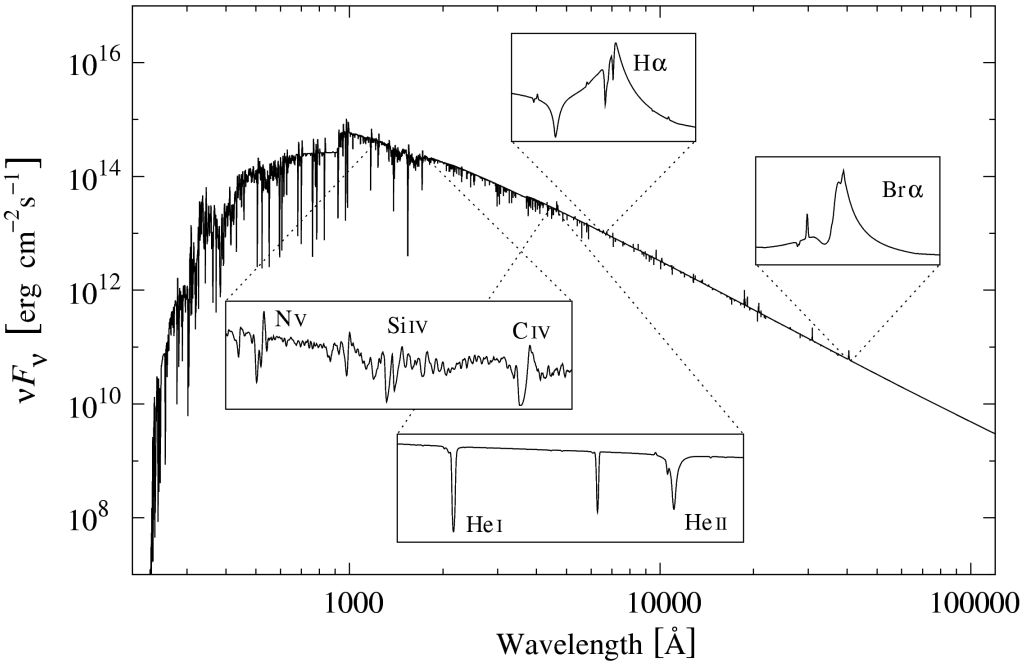


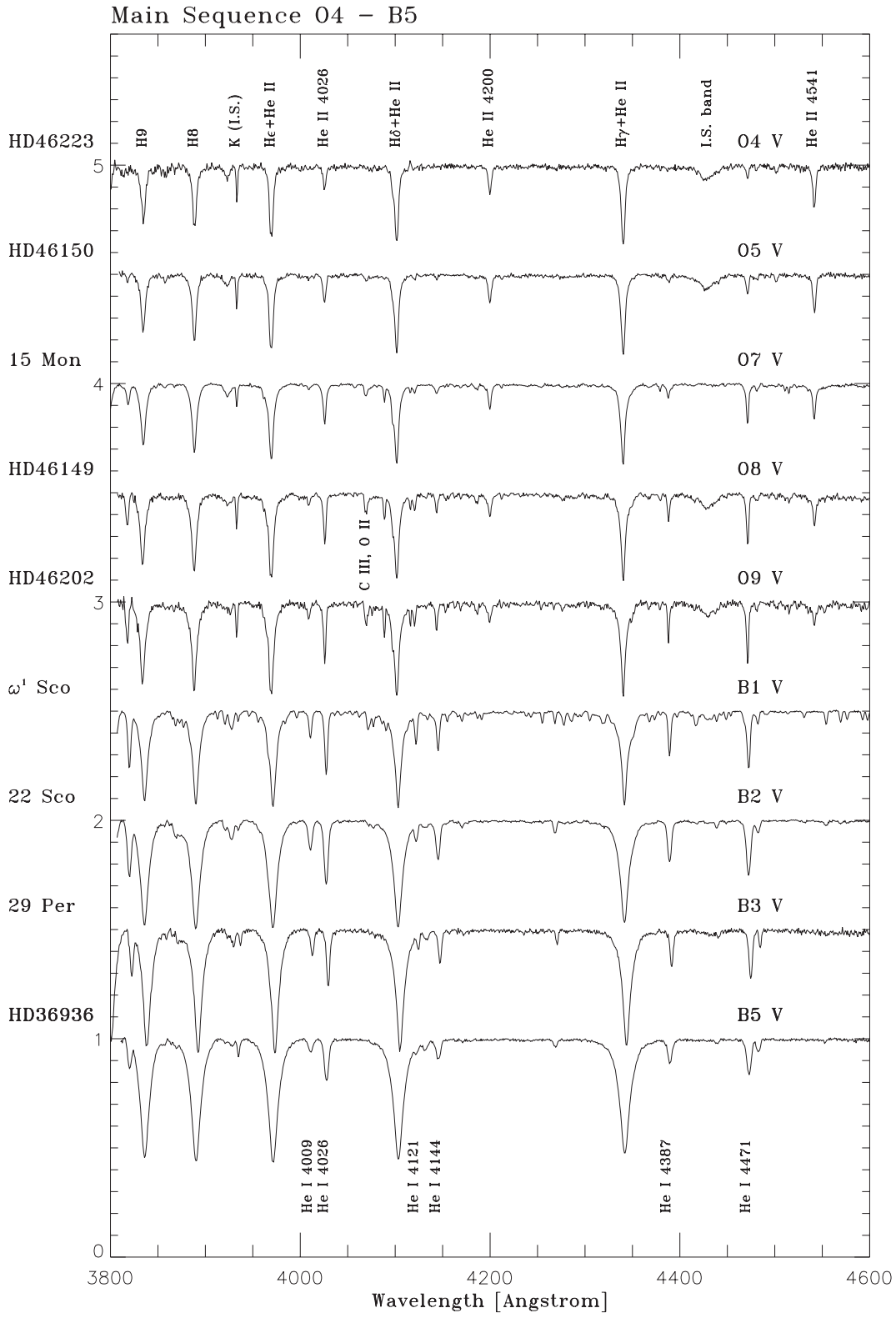
Figure 2.4: *Spectral energy distribution of an O-type star. The insets highlight specific parts of the spectrum. P Cygni type line of* N V*,* Si IV *and* C IV *are visible in the ultraviolet;* He I*,* He II *and* H I *(*Hα*) in the optical, and* H I *(*Brα*) in the near-infrared. The Lyman continuum (at* $\lambda < 912\,\text{Å}$*) shows many lines of metal ions.*

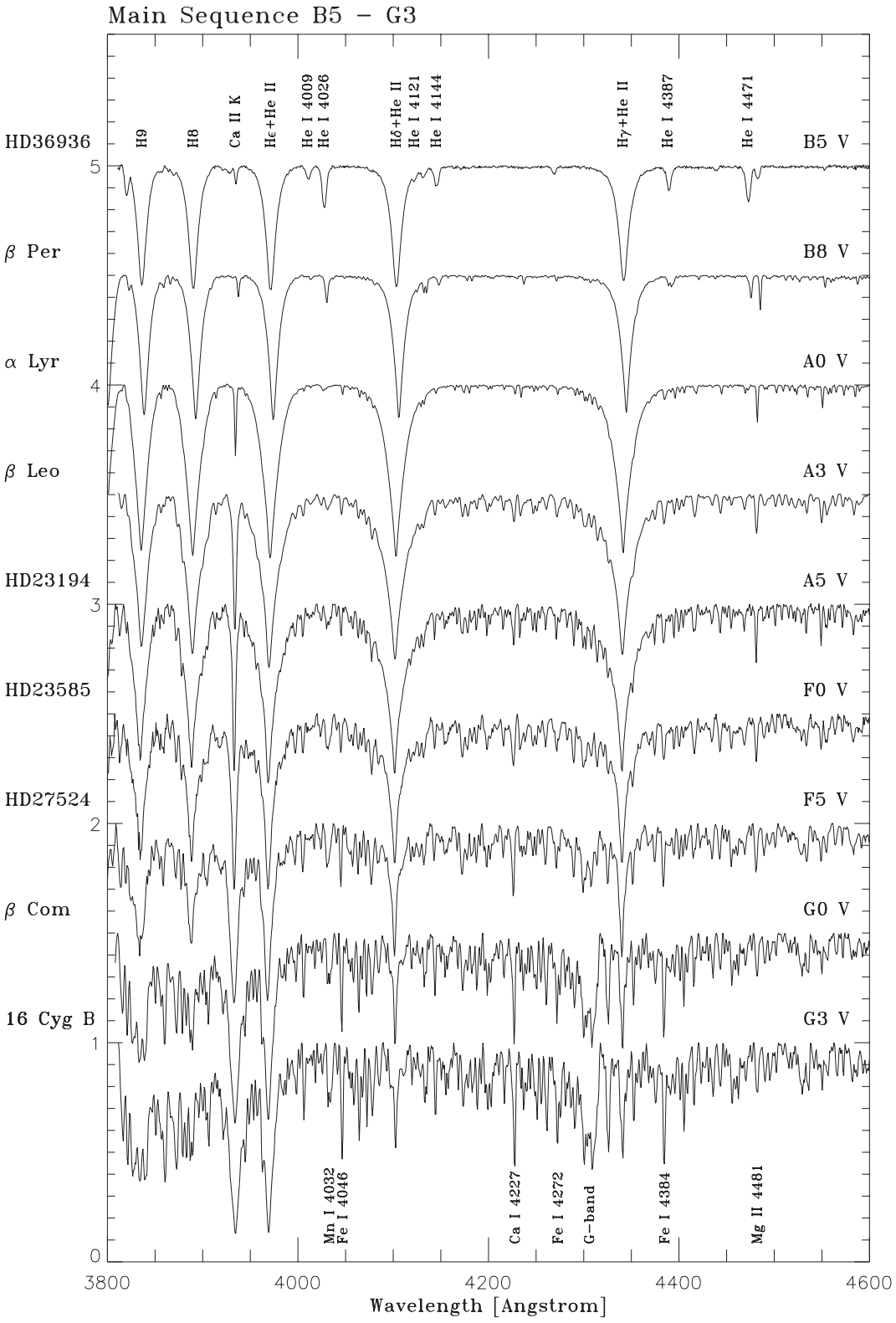
The luminosity of a star is inversely proportional to its surface gravity (see Eqs. 3.25 and 9.28 for details),

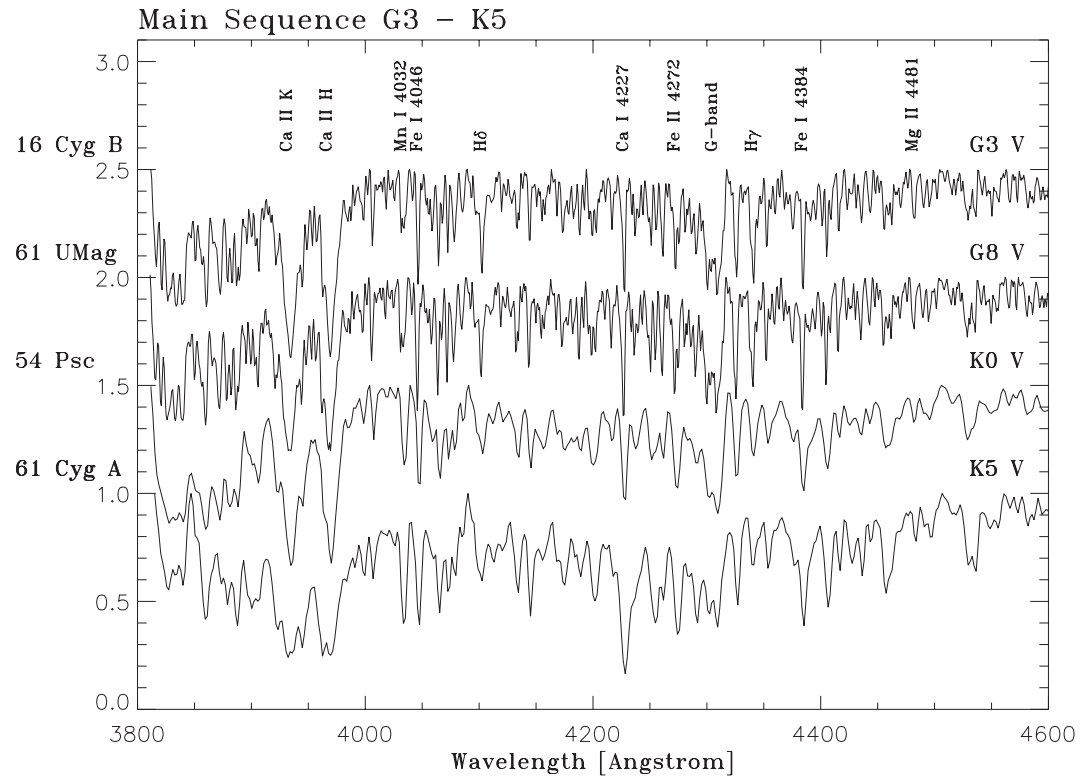
$$L = 4\pi\, R^2\, \sigma T^4 = 4\pi\, GM \frac{\sigma T^4}{g}, \tag{2.1}$$

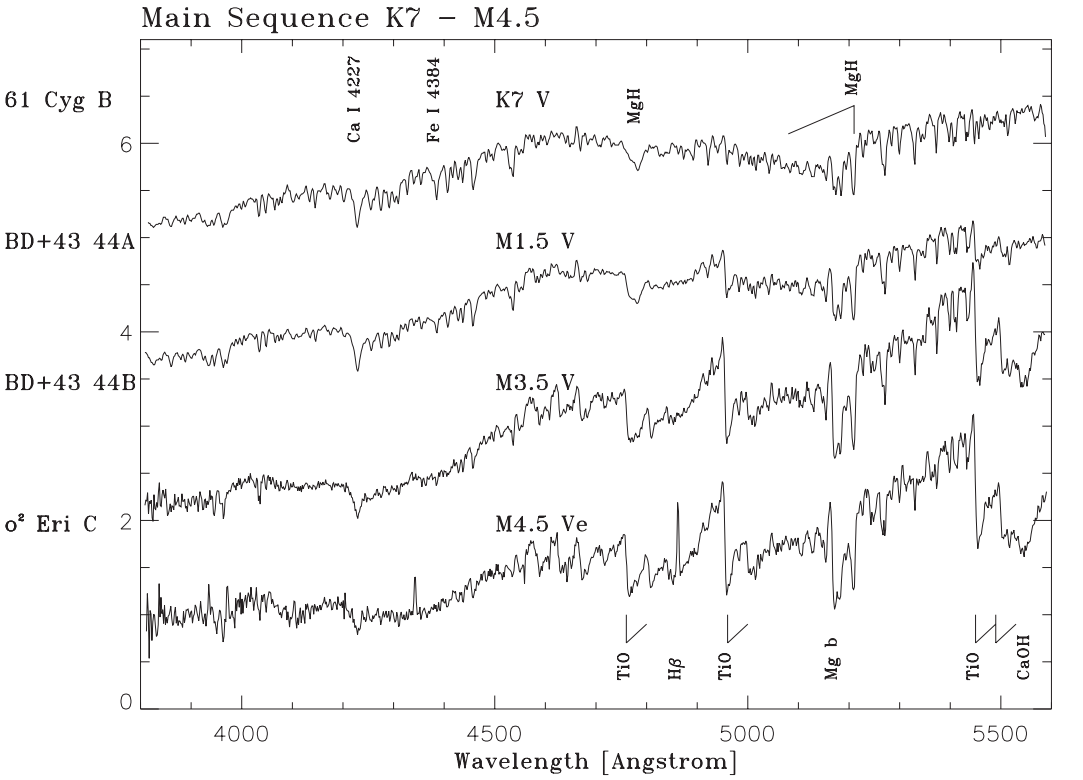
which sets the absolute scale of the pressure structure – and by that the density in the region in which the spectral lines are formed (see exercise 9.8). In early-type stars the wings of hydrogen lines are sensitive to atmospheric pressure. These lines (notably those of the Balmer series, such as Hα, Hβ, Hγ and Hδ) are therefore used to determine the luminosity class of O, B, A, and F-type stars.

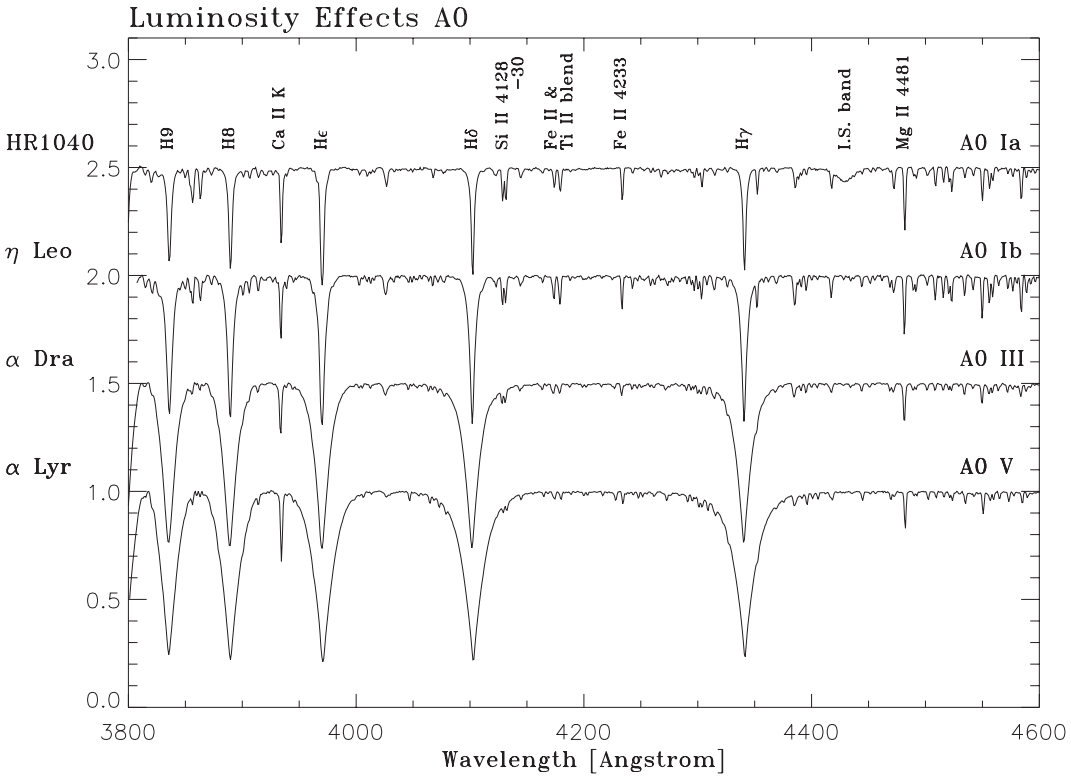
The calibration of spectral type, luminosity class and color indices (which will be discussed in § 2.6) as function of temperature and density of the medium – expressed in terms of *effective temperature* $T_{\rm eff}$ and *surface gravity* g – is done using model atmospheres. One output of such models is the spectral energy distribution or SED, of which an example is given in figure 2.4. Model atmospheres will be discussed later on in these lectures.

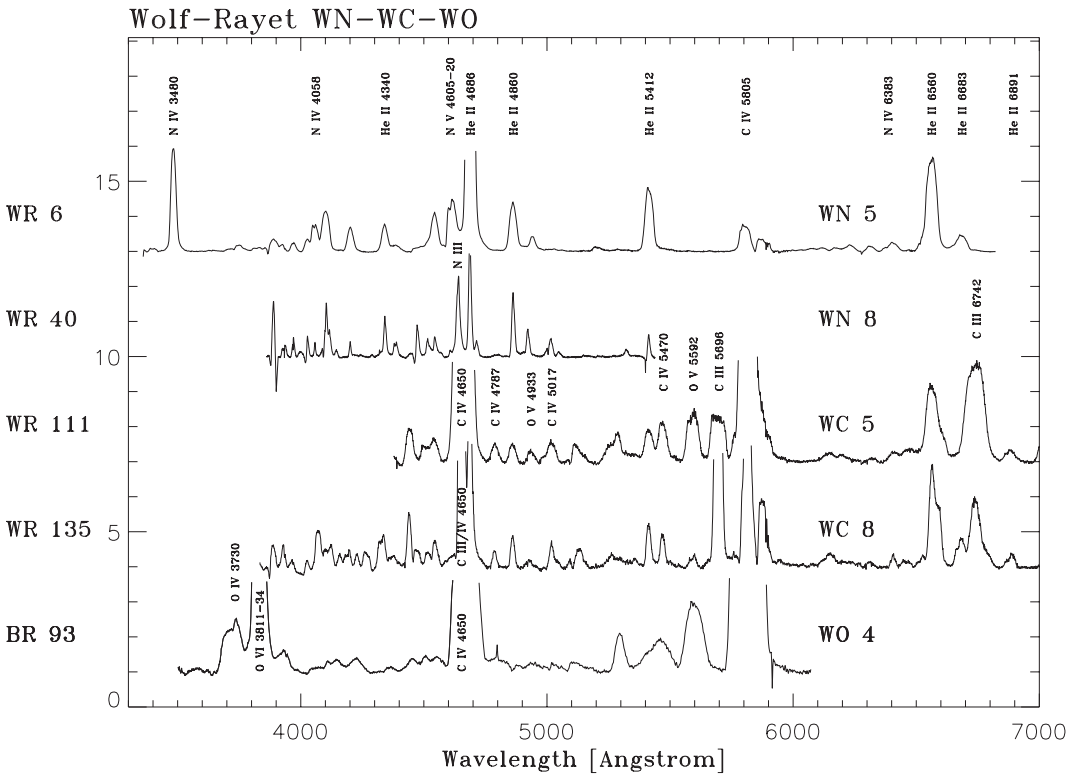

Main Sequence O4 – B5
H9
H8
K (I.S.)
Hε+He II
He II 4026
Hδ+He II
He II 4200
Hγ+He II
I.S. band
He II 4541
HD46223
O4 V
HD46150
O5 V
15 Mon
O7 V
HD46149
O8 V
C III, O II
HD46202
O9 V
ω¹ Sco
B1 V
22 Sco
B2 V
29 Per
B3 V
HD36936
B5 V
He I 4009
He I 4026
He I 4121
He I 4144
He I 4387
He I 4471
0
1
2
3
4
5
3800
4000
4200
4400
4600
Wavelength [Angstrom]


Main Sequence B5 – G3
H9
H8
Ca II K
Hε+He II
He I 4009
He I 4026
Hδ+He II
He I 4121
He I 4144
Hγ+He II
He I 4387
He I 4471
HD36936
B5 V
β Per
B8 V
α Lyr
A0 V
β Leo
A3 V
HD23194
A5 V
HD23585
F0 V
HD27524
F5 V
β Com
G0 V
16 Cyg B
G3 V
Mn I 4032
Fe I 4046
Ca I 4227
Fe I 4272
G-band
Fe I 4384
Mg II 4481
0
1
2
3
4
5
3800
4000
4200
4400
4600
Wavelength [Angstrom]


Main Sequence G3 – K5
16 Cyg B
61 UMag
54 Psc
61 Cyg A
Ca II K
Ca II H
Mn I 4032
Fe I 4046
Hδ
Ca I 4227
Fe II 4272
G-band
Hγ
Fe I 4384
Mg II 4481
G3 V
G8 V
K0 V
K5 V
Wavelength [Angstrom]


Main Sequence K7 – M4.5
61 Cyg B
BD+43 44A
BD+43 44B
o² Eri C
Ca I 4227
Fe I 4384
K7 V
MgH
MgH
M1.5 V
M3.5 V
M4.5 Ve
TiO
Hβ
TiO
Mg b
TiO
CaOH
Wavelength [Angstrom]


Luminosity Effects A0
HR1040
η Leo
α Dra
α Lyr
A0 Ia
A0 Ib
A0 III
A0 V
H9
H8
Ca II K
Hε
Hδ
Si II 4128 −30
Fe II & Ti II blend
Fe II 4233
Hγ
I.S. band
Mg II 4481
3.0
2.5
2.0
1.5
1.0
0.5
0.0
3800
4000
4200
4400
4600
Wavelength [Angstrom]


Wolf−Rayet WN−WC−WO
WR 6
WR 40
WR 111
WR 135
BR 93
WN 5
WN 8
WC 5
WC 8
WO 4
N IV 3480
N IV 4058
He II 4340
N V 4605−20
He II 4686
He II 4860
He II 5412
C IV 5805
N IV 6383
He II 6560
He II 6683
He II 6891
N III
C IV 5470
O V 5592
C III 5696
C III 6742
C IV 4650
C IV 4787
O V 4933
C IV 5017
C III/IV 4650
O IV 3730
O VI 3811−34
C IV 4650
15
10
5
0
4000
5000
6000
7000
Wavelength [Angstrom]

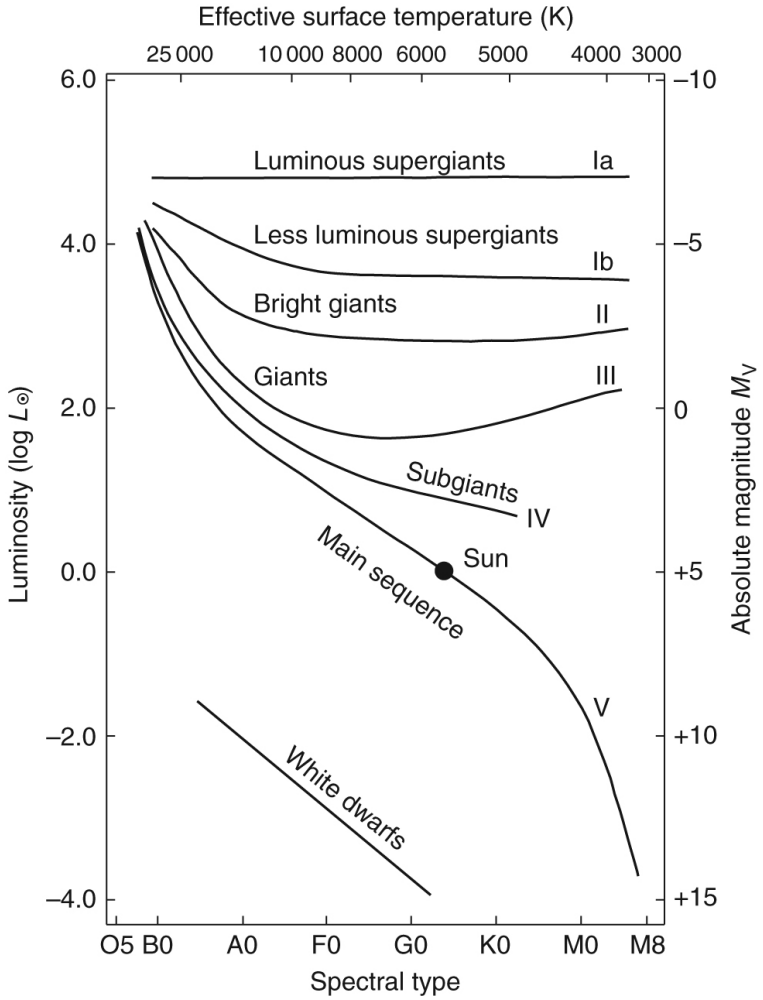


Figure 2.5: *Hertzsprung-Russell diagram showing the stellar spectral types and luminosity classes. Note the large span in luminosity of the supergiants, illustrating the need for a sub-classification in* Ia^+ *(not shown), Ia, Iab (not shown) and Ib. Source: Pradhan & Nahar,* Atomic Astrophysics and Spectroscopy.

2.2 Special stellar spectra

The vast majority of stars emit spectra that are well classified using the method we have just described. There are however stars of which the spectra show special characteristics. We will give a brief description of a few of such objects.

Be-stars

The spectra of some B- (and also O-) type stars show emission lines of, notably, the Balmer series (discrete transitions between the first and higher excited stages of neutral hydrogen). Hα is the strongest; the emission strength decreases toward higher members of the Balmer series. These stars are called B-emission (Be) stars. Be stars constitute about 15% of all B stars. In some Be stars the Be phenomenon is transient, i.e. it may be present for some time, then disappear for a while. It has been established that the rotation velocities of Be stars are on average significantly higher than that of normal B stars. Whether all B stars show the Be phenomenon during some part of their main sequence evolution or whether it is linked to binary evolution is currently being investigated. The near break-up velocities of Be stars apparently allows these stars to eject material, mainly in the equatorial regions. Line emission of this disk-like distribution of gas dominates the profile shape of the strong Balmer lines. A second characteristic of most Be stars is that they show an excess in their infrared continuum brightness.

Wolf-Rayet stars

This type of stellar spectrum was first discovered by the French astronomers Charles Wolf (1827–1918) and Georges Rayet (1839–1906), in 1867, and is characterized by broad emission lines of highly ionized elements, superposed on a very 'hot' continuum spectrum. The lines are formed in a dense, optically thick outflow or stellar wind and reflect the high mass loss of the star.

The temperatures of Wolf-Rayet (WR) stars range from that of O-type stars to well above 100 000 K. The mass-loss rates are in the order of 10^{-5} to 10^{-4} $M_{\odot}\mathrm{yr}^{-1}$. Because of the extreme character of their spectra they can not be classified using the MK scheme. They have an independent classification scheme consisting of three types. Depending of the presence of strong lines of ionized nitrogen, carbon, or oxygen one refers to them as WN-, WC- and WO stars, respectively. These are further subdivided in the spectral sequences WN2–11, WC4–9 and WO1–4, depending on the ionization-excitation conditions and the density in the outflowing atmosphere.

The *Galactic Wolf Rayet Catalogue* (v1.25, Crowther, Baker & Kus) contains 667 galactic WR stars; more than 100 are known in the Magellanic Clouds. The WR stars represent the last phase in the evolution of very massive stars (i.e. stars with an initial mass $\gtrsim 25M_{\odot}$). In absolute terms they are young stars ($\sim 5 - 10$ Myr). During the WR phase these stars end their lives in a core-collapse supernova explosion.

In a completely different category of stars WR spectra have also been observed, i.e. in some of the ($\sim$ solar mass) central stars of planetary nebulae.

Stars with P Cygni profiles

Resonance lines of abundant ions in the ultraviolet part of the spectra of O- and B-type stars often show P Cygni type profiles. Such line profiles consist of a blue shifted absorption component and a symmetric emission component. They are the result of a spherical stellar wind in which material is ejected from the star with a high velocity (on the order of 10^3 $\mathrm{km\,sec^{-1}}$). In the part of the wind that is seen in projection in front of the stellar disk, and in which the flow accelerates over a large interval in velocity (causing a shift in wavelength), the blue shifted absorption is formed. The emission originates from those parts of the outflow that are seen next to the stellar disk, and that give rise to a symmetric profile centered on the rest wavelength of the line (see chapter 17).

Carbon stars

Very cool stars have spectra that are dominated by absorption bands of molecules. They can have spectral types M, C, or S. M-type stars show strong absorption bands of TiO. C-type or carbon stars have spectra dominated by absorption bands of molecules that contain carbon. Important species are C_2, CH, and CN.

Whether a cool star has an M- or a C-type spectrum does not depend on its temperature, but on the number ratio of carbon and oxygen atoms. The reason is the exceptional stability of the CO molecule, which binds all available C and O atoms. When the star is oxygen rich, i.e. when the oxygen abundance is higher than that of carbon, a surplus of oxygen atoms will remain after the formation of carbon monoxide to form molecules such as TiO. In this case we observe an M-type spectrum. Reversely, when a star is carbon rich, a surplus of carbon atoms will remain to form carbon compounds such as the ones mentioned above. This yields a C-type spectrum.

If the C/O ratio is ~ 1 the spectral properties are intermediate between M- and C-type. These stars are classified as of S-type.

The atmospheric C/O number ratio of evolved stars of low or intermediate mass (more precisely, those that are on the asymptotic giant branch in the Hertzsprung-Russell diagram or HRD) may change over time when carbon produced in thermonuclear reactions in the stellar interior reaches the surface layers by means of convective motions. The turnover of an M-type star into a C-type star takes only a very short time ($< 10^4$ year).

Peculiar stars

The chemically anomalous or ‘peculiar’ stars form a group that in the HRD is located along the main sequence, from about spectral type B5 to F2. In total about 15% of all main sequence stars is situated in this spectral interval. They are roughly subdivided in two groups, the peculiar A (Ap) stars, and the metal line A (Am) stars. They deviate from other stars in that they

show strong lines of very specific elements. Examples are the mercury-manganese or HgMn stars and the strontium-chromium-europium or Sr-Cr-Eu stars. For these stars atmospheric abundances of the elements mentioned can be enhanced up to a factor 10^3 relative to normal values.

These deviations are thought to be the result of the exceptional stability of the atmosphere, in which a combination of gravitational forces and selective radiation pressure on spectral lines causes a differential diffusion of elements, such that some show preference to 'float' to the surface. The exceptional stability of the atmosphere is most probably linked to the absence of: *i)* a convection zone that penetrates to layers close to the surface, which explains the absence of peculiar stars later than type F2; *ii)* strong mass loss, which explains the absence of peculiar stars earlier than type B2. The atmospheric stability of peculiar stars is also illustrated by, for instance, a systematically lower X-ray luminosity of A-type main sequence stars compared to that of dwarfs of neighboring spectral types.

Magnetic fields and slow rotation also promote atmospheric stability, each process favoring specific elements to surface abundantly. Indeed, both the Ap/Bp and Am stars show much slower rotation than normal A and B stars, but stronger magnetic fields than normal A and B stars. The chemically peculiar, weakly-magnetic Am and the hotter (B-type) HgMn stars appear in binary systems more frequently than normal A and B stars. Rotational breaking due to binary tidal interactions provides a mechanism for slowing these stars sufficiently for atomic diffusion to proceed efficiently. In contrast, the strongly magnetic chemically peculiar Ap and Bp stars appear in close binary systems very rarely. Ap and Bp stars appear in widely separated binaries with roughly the same frequency as normal A and B stars, but their apparent rarity in close binary systems likely provides an important clue to the origin of their magnetic fields. A particularly interesting, although currently untested, hypothesis is that these stars are actually the product of stellar mergers. The merger process may than have generated the magnetic field.

2.3 Special binary spectra

It is estimated that about 40% of solar-type stars to 100% of massive stars are part of a binary or higher-order multiple star system. The binary fraction seems to increase with the mass of the primary (the initially most massive component of the binary systems), perhaps even reaching $\sim$100% for O star primaries. We briefly discuss the spectra of three special types of binaries.

Cataclysmic variables

Cataclysmic variables (CVs) are close binary stars in which a low mass star (of about a solar mass, which in the HRD is usually located close to the main sequence) fills its Roche lobe and

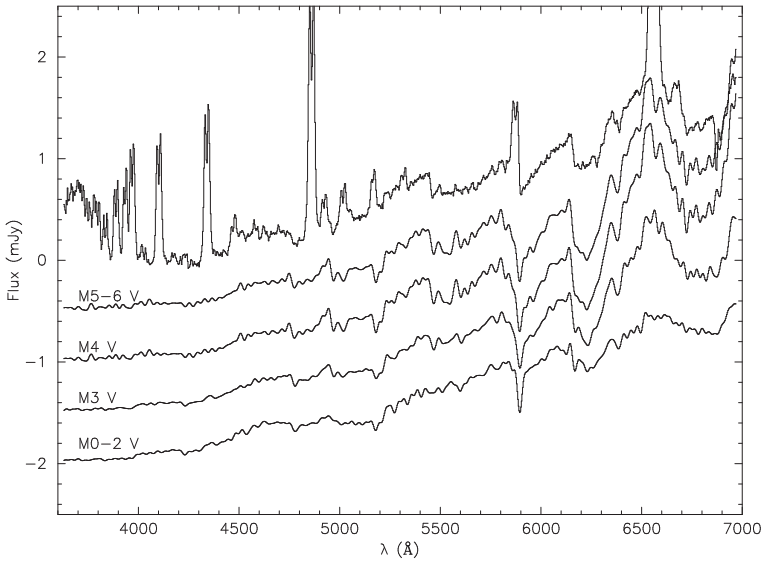


Figure 2.6: *Spectrum of the cataclysmic variable IP Pegasi (top) together with four spectra of M-type stars. IP Peg is a system with a low mass transfer rate. The spectrum shows a hybrid nature in-between that of the "blue" spectrum of the accretion disk around the white dwarf and the "red" spectrum of its M3V companion. Clearly visible in emission are the lines of the Balmer series of hydrogen and the* He I *λ5875 line. These are formed in the optically thin outer regions of the rotating disk. This can be deduced, for instance, from the fact that they are double peaked. The M3V companion reveals itself, e.g. by the strong molecular bands of TiO.*

transfers mass to a white dwarf. CVs usually have orbital periods of a few hours. The mass transfer proceeds through an accretion disk, in which material spirals in, in almost Keplerian orbits, toward the white dwarf. In most cases the disk dominates the optical light of the binary system. The spectral characteristics of CVs sensitively depend on the rate of mass transfer. If this is low (of the order of 10^{-10} $M_\odot$ per year, e.g. in dwarf novae in quiescence) one sees strong emission lines, notably those of the Balmer series (see figure 2.6), superimposed on a blue continuum. These lines originate from a large optically thin regime in the outer parts of the disk. When the mass transfer rate is high (of the order of 10^{-8} $M_\odot$ per year) the continuum emission increases relative to that of the lines (because also the outer parts of the disk become optically thick). In some cases the lines disappear (e.g. during a dwarf nova eruption).

Low Mass X-ray Binaries and High Mass X-ray Binaries

Low Mass X-ray Binaries (LMXBs) are similar to CVs, except that the compact object is a neutron star or a black hole i.s.o. a white dwarf. The optical spectra of LMXBs show a blue continuum (corresponding to a temperature of 30 000 K or higher), on which superposed are a number of weak emission lines. The most important of these lines are Balmer series transitions, He II λ4686, and C III/N III λ4640-50. The spectra originate in the accretion disk around the compact star, where x-ray radiation from the compact object is being converted in

optical and ultraviolet photons (so called x-ray heating or "reprocessing").

In High Mass X-ray Binaries (HMXBs) the star accompanying the neutron star or black hole is a massive star (typically $> 10\ M_\odot$) of spectral type O or early-B. Accretion of matter on the compact object can be different for LMXBs and HMXBs. In the case of LMXBs, the small and low mass companion star fills and overflows its Roche lobe, therefore accretion of matter always occurs through the formation of an accretion disk. In the case of HMXBs, accretion can also occur through an accretion disk, for systems in which the companion (supergiant) star overflows its Roche lobe. There are however two alternatives. The first concerns dwarf to giant Be-stars, i.e. stars with a circum-stellar disk (see above). Here it is when the compact object – on a wide and eccentric orbit – crosses this disk, that accretion periodically occurs. These systems are sometimes referred to as Be-HMXBs. The second case is when the massive star (often a supergiant) ejects a dense radial stellar wind (see Chapter 17), and the compact object directly accretes the stellar wind through Bondy-Hoyle-Littleton processes. These are sometimes referred to as Sg-HMXBs.

We point out that there also exist X-ray binary systems for which the normal star if of intermediate mass (typically between 1 and 10 $M_\odot$), called IMXBs for Intermediate Mass X-Ray Binaries.

2.4 Brown dwarfs and planets

The masses of brown dwarfs are too low to allow thermonuclear burning of hydrogen in their cores, resulting in a steady decline in both luminosity and effective temperature with time[5]. Deuterium is fused during the contraction phase. Lithium however always remains present (and it is therefore the detection of Li that is used to identify these dim objects as brown dwarfs). Very low mass stars and brown dwarfs are probably the most common objects in our Milky Way. The warmest brown dwarfs have M-type spectra and are dominated by molecular absorption bands of H_2, CO, H_2O and metal oxides such as TiO and VO (see Fig. 2.7).

The cooler L-type brown dwarfs (with effective temperatures $T_{\rm eff} \lesssim 2\,200$ K) are characterized by strong metal-hydride (FeH, CrH), H_2O and CO molecular absorption; and alkali lines, including the heavily pressure-broadened Na I and K I doublets. L dwarfs also show evidence for condensed clouds in their photospheres, which give rise to highly reddened spectral energy distributions and absorption features from silicate grains. Ti and V are locked up in solid state particles causing the TiO and VO molecular bands to disappear.

Even cooler Jupiter-like objects with masses less than $\simeq 13\ M_{\rm J}$ do not even burn deuterium. If the spectra of such objects show strong CH_4 bands then $T_{\rm eff} \lesssim 1\,300$ K and spectral type T is used to characterise them. Other typical features are H_2O, NH_3, and strong collision-induced

[5]The mass limit for sustained hydrogen fusion is roughly 0.072 $M_\odot$ (71 Jupiter masses) for a Solar metallicity gas mixture, increasing to 0.090 $M_\odot$ for a pure hydrogen gas (e.g. Chabrier & Baraffe 2000).

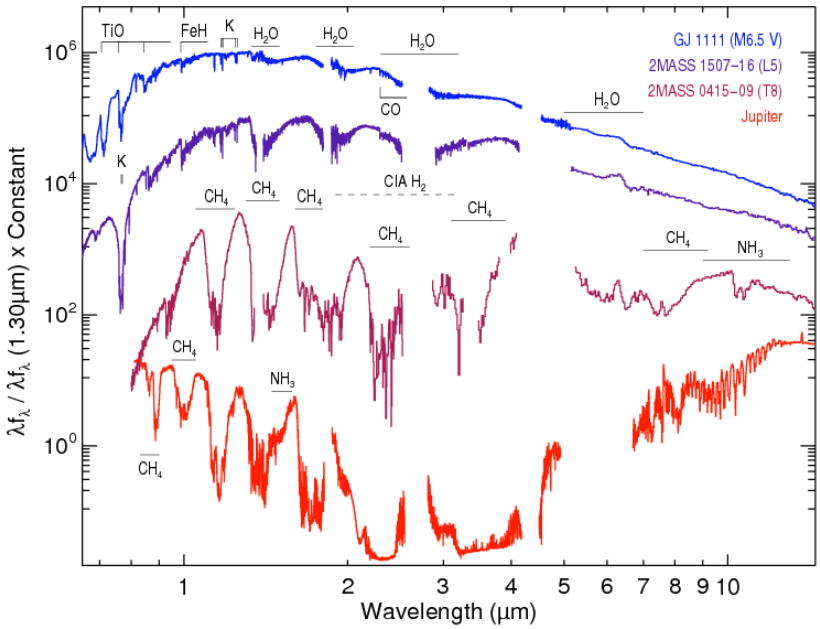


Figure 2.7: *Spectra of representative M-type, L-type, and T-type dwarfs, compared to data for Jupiter. The spectra are arbitrarily normalised. Major molecular absorption bands characterising these spectra are labeled. Atomic K I absorption is also labeled, which produces a substantial pressure-broadened line feature spanning 0.7-0.85 μm in L and T dwarf spectra. Note that Jupiter's emission shortward of $\sim$4 μm is dominated by scattered solar light modulated by CH_4 and NH_3 absorption features, while the dwarf spectra are entirely due to emergent flux From: Burgasser 2009*

H_2 absorption.

A formal and robust definition of the spectral type Y, for even cooler objects, is not yet in place. The appearance of an absorption feature of potentially ammonia (NH_3) at 1.55 μm has been proposed as a classifier. Objects that show this feature have been modelled to have $T_{\rm eff} \lesssim$ 400 K.

Brown dwarfs differ from Jupiter-like objects, as are presently being discovered around nearby stars, in their genesis. For example, the bulk of the discovered L dwarfs is 'free floating', i.e. they exist as individual objects in space, contrary to (exo)-planets. The classification of exoplanet spectra is complicated by the presence of the much brighter host star. A promising technique is that of transit spectroscopy, in which the extinction of background host-starlight passing through the exoplanet atmosphere is used to probe the composition and vertical height of the exoplanet atmosphere. Figure 2.8 shows such a classification for hot Jupiters in the (equilibrium) temperature range of 1000-2500 K.

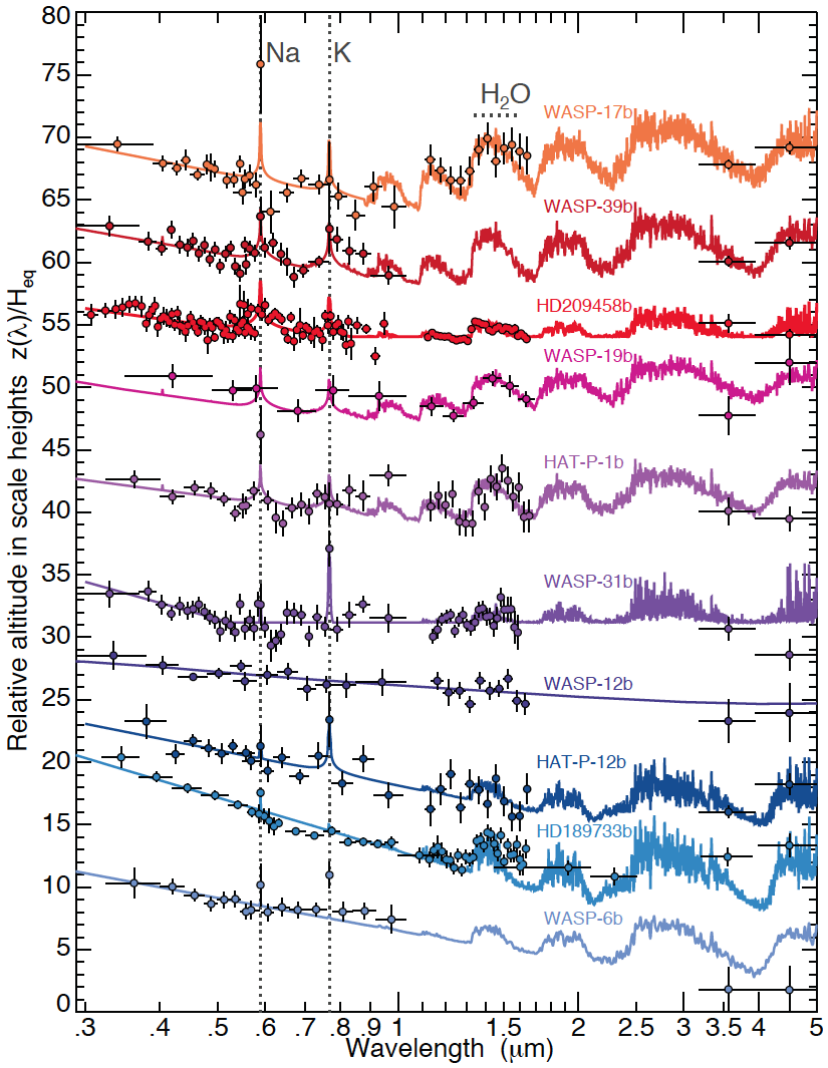


Figure 2.8: *Spectral sequence of hot-Jupiter exoplanets. The dots are transmission observations; the solid coloured lines show fitted atmospheric models. Horizontal and vertical error bars indicate the wavelength spectral bin and 1σ measurement uncertainties, respectively. Prominent spectral features of Na, K, and H_2O are indicated. Plotted on the vertical axis is the transit-measured altitude $z(\lambda)$ of a hydrostatic atmosphere normalised to the pressure scale height $H_{\rm eq}$, which is a measure of the degree of cloudiness (see Chapter 12). The planet with smallest $z(\lambda)/H_{\rm eq}$ is at the top, the one with largest ratio at the bottom. Planets with predominantly clear atmospheres (top) show prominent alkali and H_2O absorption; heavily hazy and cloudy planets (bottom) have narrow alkali lines and H_2O absorption that is partially or completely obscured. The latter planets also feature the strongest optical (0.3 - 0.57 μm) to infrared (3 - 5 μm) continuum slopes. The relatively weak H_2O features in the hazy and cloudy planets is interpreted as clouds obscuring the deeper layers where water emission originates. From: Sing et al. 2016*

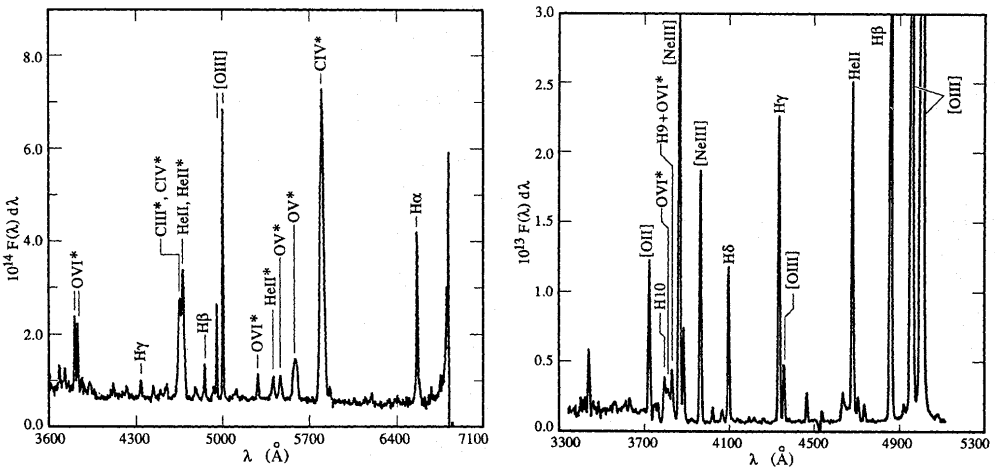


Figure 2.9: *Spectra of the Planetary nebulae NGC 1501 (left) and IC 1297 (right) in visible light. Planetary nebulae are characterized by strong emission lines and a (very) weak continuum. The strongest nebular lines in PN spectra are often those of the forbidden [O III] λ4959 and 5007 transitions. The lines that have the superscript * are from the central star, which in both cases has an effective temperature of $T_{\rm eff} \sim 80\,000$ K; the nebulae have a electron temperature $T_{\rm e} \sim 10\,000$ K. Source: Stanghellini, Kaler, & Shaw 1994, A&A 291, 604.*

2.5 Nebulae

Planetary nebulae

Diffuse objects that can not be resolved in stars are referred to as nebulae. It may be clear that in many cases this definition says more about the spatial resolution of the instrument that is used than about the nature of the observed object. Planetary nebulae were termed as such by Frederick William Herschel (1738–1822), who saw them with his telescope and thought they looked somewhat like the disks of planets. The first spectrum of a planetary nebula was recorded in 1864 by William Huggins (1824–1910). Huggins discovered that a single emission line seemed responsible for the light. At higher spectral resolution this single line could be resolved in three individual lines, of which Hβ is one.

Planetary nebula contain a central star. The nebula is ejected by this star. Planetary nebula represent a short (of the order of 10^4 year) period near the end of the lives of low mass stars, after evolving of the asymptotic giant branch but before reaching the white dwarf stage. The spectra of planetary nebulae are dominated by line emission; continuum emission is relatively weak. The hydrogen line emission in the nebula results from the illumination by Lyman continuum photons from the very hot (30 000 to $\sim$ 150 000 K) central star, which ionises the nebular gas. After recombination a line cascade occurs (e.g. in the Balmer series) giving rise to the emission lines. Planetary nebulae also show forbidden line emission, e.g. [O III] λ4959

and 5007. This radiation originates from collisional excitation of metastable levels, which as a result of the extremely low densities in nebulae are no longer depopulated by collisional processes. Figure 2.9 shows two examples. We will discuss the line formation in gaseous nebulae at a conceptual level in more detail in chapter 15. Further quantitative elaborations will be given in chapter 18.

Nebulae around massive stars

Massive stars may also be surrounded by nebulae. These may form from material ejected by the star, typically in evolutionary phases characterised by a high mass loss in a stellar wind. Nebulae are observed around for instance Luminous Blue Variable stars, Wolf-Rayet stars and red supergiants. The nebular spectra of LBVs en WRs are dominated by recombination lines, just as in the case of planetary nebulae. In the infrared one often observes the thermal emission of solid state material (or dust), that may condense in the stellar outflow or in the nebular environment.

2.6 The continuum energy distribution

Connected to the changes in relative line strengths that are observed along the spectral sequence, and which are caused by changes in the ionisation and excitation conditions in the stellar atmospheres, is a change in the continuum energy distribution of the emitted light.

Libraries of optical and near-IR spectra illustrate the changes in the continuum energy distribution as a function of spectral type. An example is the VLT/UVES Paranal Observatory Project by Bagnulo et al. (2003). The spectra cover the wavelength range from 3040 to 10400 Å (save for a few narrow wavelength gaps) and have a spectral resolving power (see eq. 13.3) of about 80 000. For most of the spectra, the typical signal-to-noise ratio in the V band is between 300 and 500. All this implies that both the individual spectral lines and continuum energy distribution are clearly visible. The spectra can be obtained and/or visualised using the UVES POP Archives. A second example is the X-Shooter Spectral Library, covering the wavelength range 3000 to 25000 Å (save for a few gaps) at a spectral resolution $R \sim 10,000$ (Verro et al. 2022). This library contains spectra for 683 stars of spectral types O through M, as well as AGB-stars. Both original and galactic dust corrected spectra are available.

Changes in the spectral energy distribution can be measured by one or more *color indices*. The best known are the $B-V$ and $U-B$ indices introduced by Johnson & Morgan (1953). The central wavelengths of U, B, and V broad-band filters, as defined by Johnson, are at 3600, 4400, and 5500 Å, respectively. The typical filter width is about 1000 Å. Their transmission efficiency (convolved with the detector transmission) as a function of wavelength is given in figure 2.10 for different realisations (see below).

It is customary to define the index in such a way that the magnitude in the long wavelength

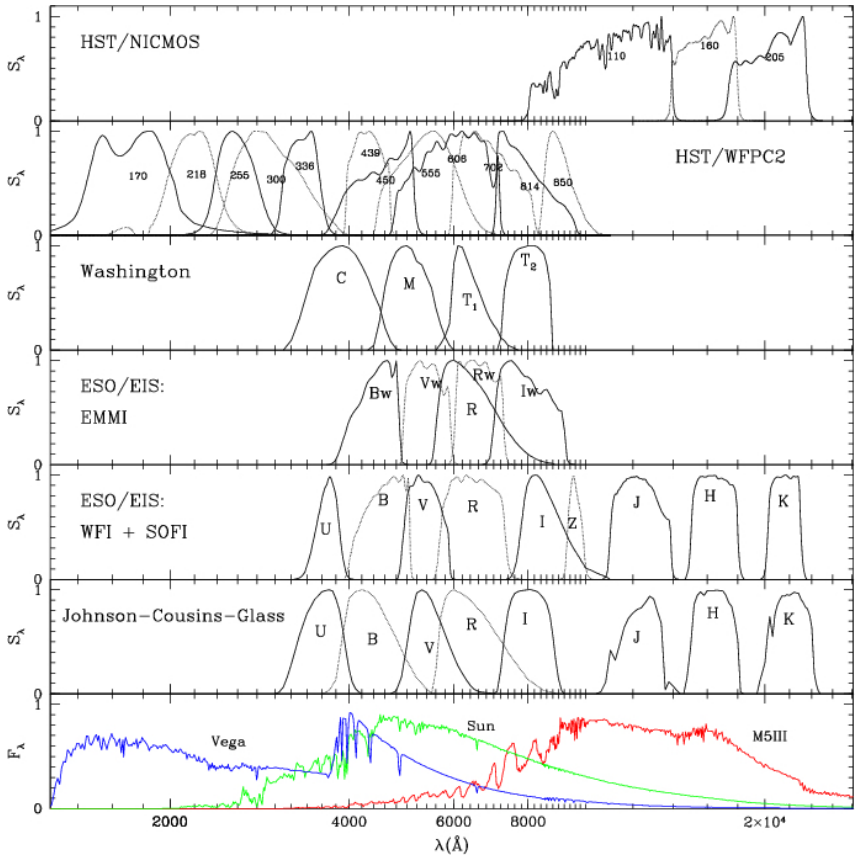


Figure 2.10: *Filter transmission curves convolved with detector transmission curves for six photometric systems: (1) HST/NICMOS, (2) HST/WFPC2, (3) Washington, (4) ESO/EMMI, (5) ESO/WFI UBVRIZ + ESO/SOFI JHK, and (6) Johnson-Cousins-Glass. To facilitate a relative comparison, the curves have been re-normalized to their maximum value of S_λ. The bottom panel presents the spectra of Vega (A0 V), the Sun (G2 V), and an M5 giant, in arbitrary flux scales. From: Girardi et al. (2002).*

filter is subtracted from that in the short wavelength filter, such that an increase in the color index signifies a 'redder' spectrum. The run of the $B-V$ and $U-B$ indices (in the Johnson system) with spectral type and luminosity class is shown in figure 2.11 and in table B.5.

The $B-V$ color index is a relative measure of the temperature of the star in the Paschen continuum (see figure 6.4); the $U-B$ color index too is a measure of temperature. In addition it probes the extend of the Balmer jump, which not only depends on temperature but also somewhat on density.

If one plans to use these indices to determine the spectral type and luminosity class of a star it is important to realize that extinction in the interstellar medium – as a result of the interaction of starlight with 'interstellar dust' – can affect the intrinsic continuum energy distribution (see chapter 19). The importance of extinction increases with decreasing wavelength and therefore

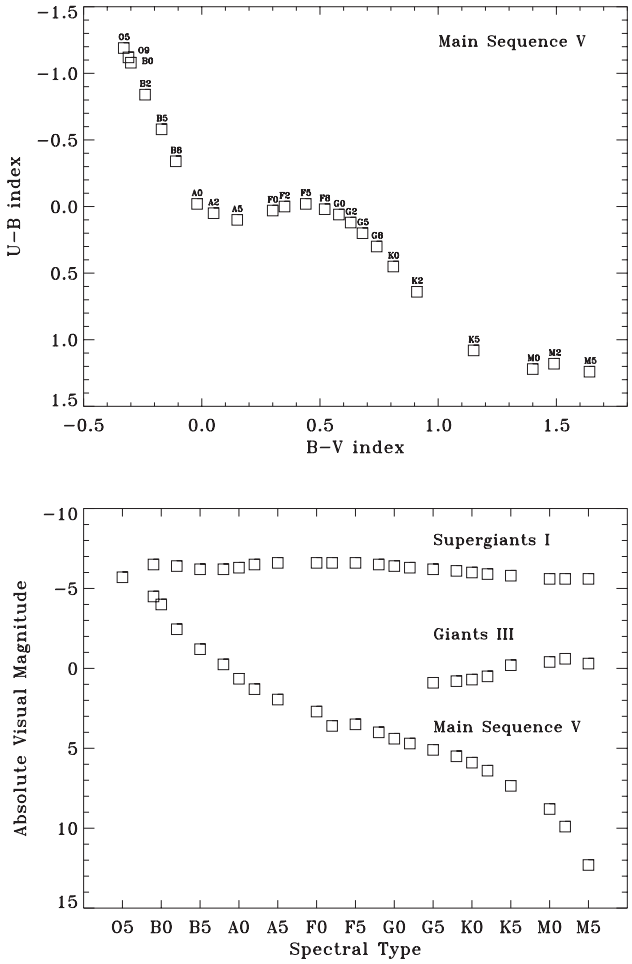


Figure 2.11: Top: *The color indices $U-B$ and $B-V$ as a function of spectral type and luminosity class.* Bottom: *The absolute visual magnitude versus spectral type, i.e. the Hertzsprung-Russell diagram.*

also has an effect on the color indices. Interstellar extinction is typically important for stars at relatively large distances in the plane of the Milky Way and for stars in star-forming regions. For nearby stars (say $d \lesssim 100$ pc) interstellar extinction is negligible. The shape of spectral lines is not affected by interstellar extinction as they cover only a very small wavelength band, such that one may ignore the wavelength dependence of the extinction. Often stellar spectra show absorption lines that originate from extinction of starlight passing through interstellar gas.

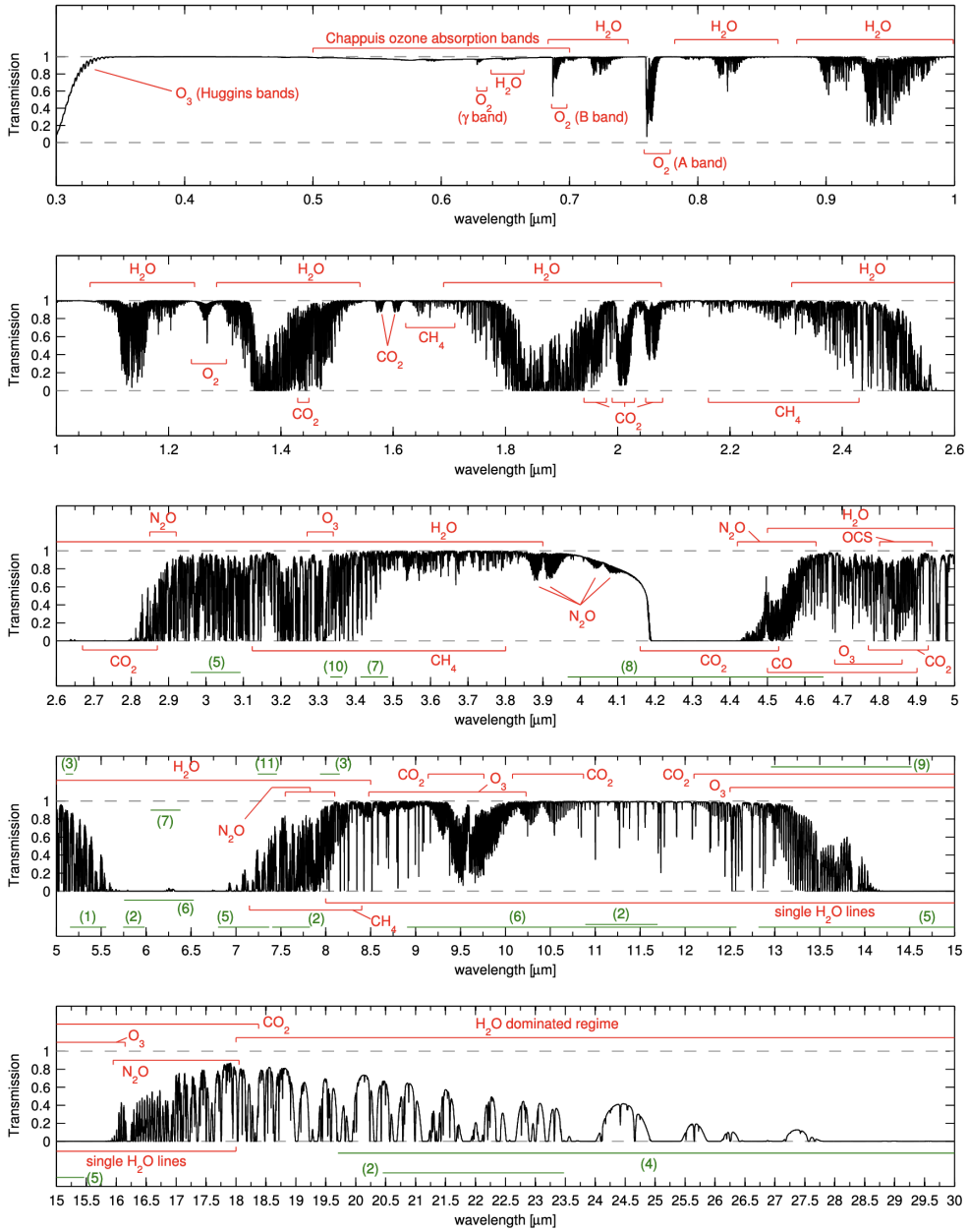


Figure 2.12: *Synthetic absorption spectrum of the sky between 0.3 and 30 μm at a spectral resolution of $R \sim 10\,000$ using the annual mean profile for Cerro Paranal (Noll et al. 2012). The eight molecules O_2, O_3. H_2O, CO, CO_2, CH_4, OCS, and N_2O are important contributors to the telluric absorption. The red regions mark the ranges where they strongly impact the transmission. Green regions denote (minor) contributions from (1) NO, (2) HNO_3, (3) COF_2, (4) H_2O_2, (5) HCN, (6) NH_3, (7) NO_2, (8) N_2, C_2H_2, (10) C_2H_6, and (11) SO_2. From: Smette et al. 2015.*

	Paranal (FORS1 Bessell)					La Silla (COMBO17)					Mercator (MAIA)			
	$\lambda_{\rm m}$	$\lambda_{\rm min}$	$\lambda_{\rm max}$	$\log C_m$		$\lambda_{\rm m}$	$\lambda_{\rm min}$	$\lambda_{\rm max}$	$\log C_m$		$\lambda_{\rm m}$	$\lambda_{\rm min}$	$\lambda_{\rm max}$	$\log C_m$
U	0.36	0.32	0.40	-8.45	U	0.37	0.36	0.40	-8.41	u	0.36	0.32	0.39	-8.42
B	0.43	0.35	0.51	-8.19	B	0.46	0.39	0.51	-8.23	g	0.49	0.40	0.56	-8.29
V	0.55	0.48	0.68	-8.45	V	0.54	0.48	0.60	-8.42					
R	0.67	0.56	0.93	-8.68	R	0.65	0.56	0.77	-8.66	r	0.64	0.54	0.71	-8.64
I	0.79	0.70	0.89	-8.93	I	0.86	0.74	0.98	-9.03					
	Paranal (NACO)					2MASS								
J	1.26	1.06	1.44	-9.54	J	1.24	1.08	1.41	-9.50					
H	1.66	1.44	1.88	-9.95	H	1.66	1.48	1.82	-9.95					
K_s	2.14	1.90	2.42	-10.36	K_s	2.16	1.95	2.36	-10.37					
L'	3.81	3.40	4.27	-11.30										
M'	4.78	4.39	5.21	-11.69										
	Paranal (MIDI)													
N8.7	8.71	7.59	9.74	-12.70										
N11.3	11.3	10.61	12.07	-13.15										

Table 2.2: *Properties of the optical/infrared filters of six common system. The mean wavelength $\lambda_\circ$, minumum wavelength $\lambda_{\rm min}$, and maximum wavelenth $\lambda_{\rm max}$ are in micron. The calibration constant C_m of filter m has dimensions erg cm^{-2} sec^{-1} Å^{-1}. From: SVO Filter Profile Service.*

The jungle of photometric systems

For different reasons no universal photometric system exists. Johnson (1966, ARA&A 4, p193), who measured the photometric magnitudes of a large number of stars in ten different filters, lost his filter set before he could measure the filter transmissions. Later, others tried to design filters with properties that best recovered the colors of the stars measured by Johnson – but without getting a perfect match. One should also realize that the quantum efficiency of the detector and the optical efficiency of the lenses in the telescope may differ for each observatory, as, obviously, will the atmospheric transmission. Especially in the near- and mid-infrared the atmospheric transmission is a problem. Figure 2.12 shows the synthetically computed atmospheric transmission at Cerro Paranal, where the VLT telescopes are located. Note that the near-infrared filters are designed to make optimal use of relatively transparent 'windows' in the earth atmosphere.

The measured flux in filter m is related to the true flux distribution of the star at the distance to the earth, $\mathcal{F}_\lambda(d)$, where d is the distance, as

$$\mathcal{F}_m(d) = \frac{\int_0^\infty \mathcal{F}_\lambda(d)\mathcal{S}_m(\lambda)\,d\lambda}{\int_0^\infty \mathcal{S}_m(\lambda)\,d\lambda}, \tag{2.2}$$

where $\mathcal{S}_m$ = (filter transmission) $\otimes$ (detector quantum efficiency) $\otimes$ (optical efficiency) $\otimes$ (atmospheric transmission) is the *response function*. Defining $\int_0^\infty \mathcal{S}_m(\lambda)\,d\lambda \equiv \mathcal{S}_m^\circ$, yields for the apparent magnitude

$$m = -2.5\log\left[\frac{1}{S_m^\circ}\int_0^\infty \mathcal{F}_\lambda(d)\,\mathcal{S}_m(\lambda)\,d\lambda\right] + 2.5\,C_m. \tag{2.3}$$

The *filter calibration constant* may be defined such that $m = 0$ for the Vega (α Lyrae) system. If the spectral energy distribution of Vega at earth is $Vg_\lambda(d)$ it follows that

$$C_m = \frac{1}{S_m^\circ}\int_0^\infty Vg_\lambda(d)\,\mathcal{S}_m(\lambda)\,d\lambda. \tag{2.4}$$

Vega is the fifth brightest star in the night sky. Though ideally one would have wished Vega to be a star of constant brightness, it turns out that it varies slightly (0.09 mag in V). For this and other reasons, the calibration constants are often derived on the basis of a set of standard stars.

Filter properties and calibration constants C_m for a few often used filter systems are given in table 2.2. An overview of hundreds of filter systems is provided at the Filter Profile Service. Using C_m one may obtain the measured flux, $\mathcal{F}_m$, corresponding to the observed or *apparent magnitude* m, from

$$\log \mathcal{F}_m = -0.4\, m + C_m. \tag{2.5}$$

As the tabulation of C_m uses units erg cm^{-2} sec^{-1} Å^{-1}, this too is the unit for the measured flux. If, for instance, we would be considering the Paranal/FORS1 Bessell V filter, the above equation would tell us that for a star with $V = 0$ the measured flux is $10^{-8.45} = 3.55 \times 10^{-9}$ erg cm^{-2} sec^{-1} Å^{-1}. For a star with $V = 5$, the measured flux is 100 times less.

Exercise 2.1

We have classified the spectrum of a star as M2 III and have measured its visual magnitude: $V = 1.63$ magn. We assume that in the line-of-sight towards the star there is no interstellar extinction at visual wavelengths: $A_V = 0$ magn. Use the equation for the *distance modulus*

$$m_f - M_f = 5 \log d - 5 + A_f \tag{2.6}$$

where m is the apparent magnitude and M the absolute magnitude of photometric filter f, d the distance in parsec, and A_f the interstellar extinction in filter f (see also eq. 19.2), as well as table B.5 to determine its distance. This procedure of distance determination is known as the spectroscopic parallax method.

Exercise 2.2

A large survey provides a catalogue of all sources brighter than V = 20 mag in a certain part of the sky. We are interested in an M5 V star at a distance of 400 pc that is in the observed field. Is this star listed in the catalogue? Use table B.5 for a measure of the absolute magnitude of the star.

Exercise 2.3

Consider the two stars whose properties have been described below. The symbols in the table have their usual meaning.

Star	V	$B-V$	M_V	$T_{\rm eff}$	Spectral Class	BC
Betelgeuse	0.42	1.84	-5.47	3370	M2 Ib	-1.62
Gliese 887	7.34	1.48	9.75	3520	M2 V	-1.89

How much larger in radius is Betelgeuse than Giese 887?

Exercise 2.4

The top panel of Figure 2.11 shows a plot of the color-color diagram in which $U-B$ is plotted as a function $B-V$. Explain why the $U-B$ color shows a local maximum at about spectral type F5.

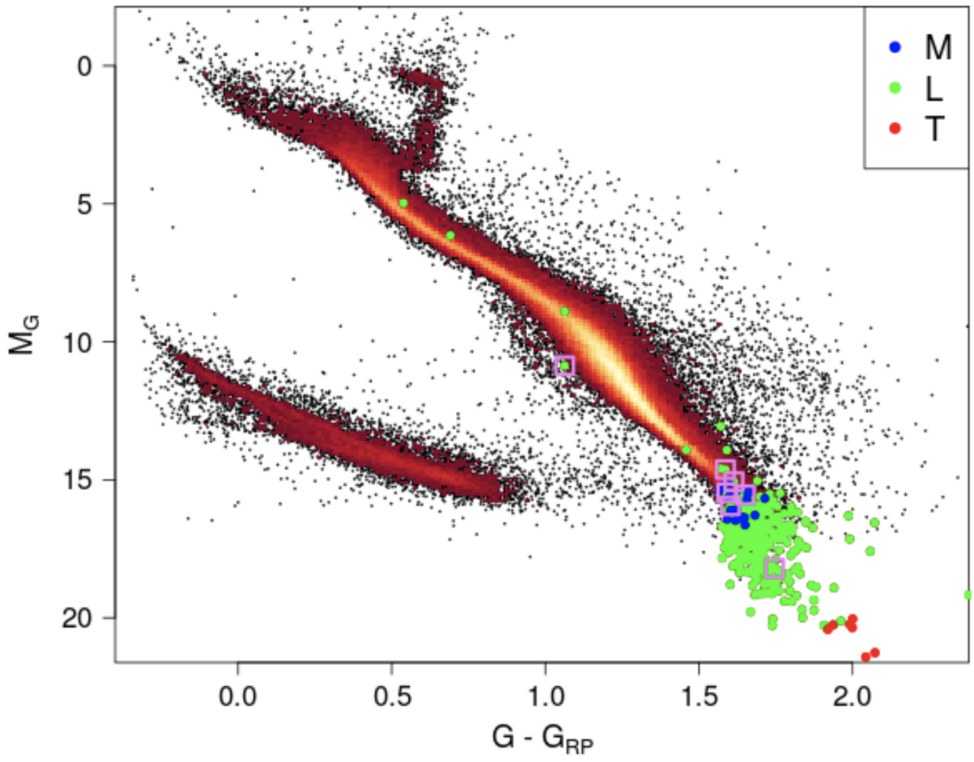


Figure 2.13: *HRD of 240 703 stars with parallax $\varpi > 10$ (or distance $d < 100$ pc) observed with GAIA. Cross matches with the* Gaia Ultracool Dwarf Sample *(GUCDS; Smart et al. 2017) of brown dwarfs and gas giants are given the color blue (M-type object), green (L type) or red (T type). The magnitudes of (in any case) four green objects are too bright, most probably because of a cross-match issue. This leaves 21, 443, and 7 of M-, L, and T-type objects, respectively. Pink squares are added around stars with tangential velocity $v_T > 200\,\mathrm{km\,sec^{-1}}$. From: Gaia Collaboration 2018).*

Exercise 2.5

Figure 2.13 shows the HRD diagram of 240 703 sources within 100 pc. One clearly recognizes the main sequence and the giant branch (starting at $M_G \sim 3$ and $G - G_{RP} \sim 0.6$). Labelled with blue (M-type objects; 21 in number), green (L type; 443), and red (T type; 7) are objects that have been identified as of sub-stellar mass.

a) Is it likely that the relative numbers of M-, L-, and T-type objects of sub-stellar mass represent the intrinsic relative numbers of these types of objects in the solar neighborhood?

b) The white dwarfs reside on the branch to the lower left of the main sequence. Notice that at $G - G_{RP} > 0.9$ this branch seems to bridge to the regime of the low-mass stars and sub-solar mass objects. What could be the nature of the objects located in this 'bridge'?

c) Say, we want to find sub-solar mass objects using photometry, e.g. by focussing on

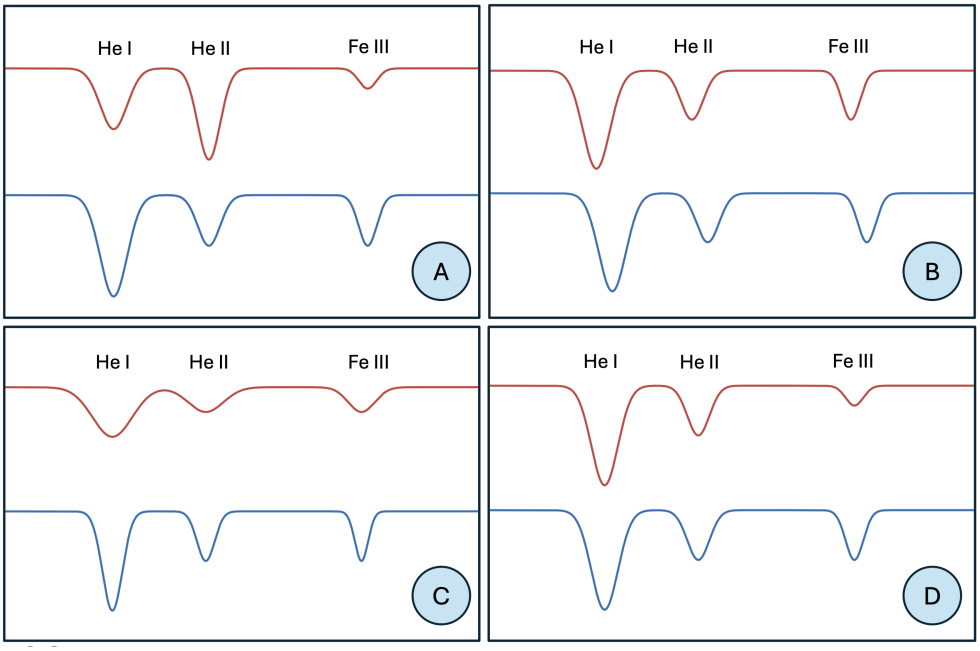


Figure 2.14: *Panels of the same small (hypothetical) spectral window for two different stars observed with the same spectrograph (settings) and telescope.*

objects with $M_{\rm G} \gtrsim 15.5$ and $G - G_{\rm RP} \gtrsim 1.6$. Would we pick up only sub-solar mass objects? If not, how could we further eliminate sources of a different nature?

d) Notice that quite a number of black dots are to the right of the lower part of the main sequence (at $M_{\rm G} \sim 10 - 15$ and $G - G_{\rm RP} \gtrsim 1.5$). What could be the nature of these objects?

Exercise 2.6

In Fig. 2.14, each panel shows the same small (hypothetical) spectral window for two different stars observed with the same spectrograph (and spectrograph settings) and telescope. The spectra are normalized to the continuum and have the same flux (y-axis) and wavelength (x-axis) scale. For convenience one of the spectra is offset in the y-direction. In each case, one physical property is different for the two stars. Identify, per panel, which physical property is different and briefly explain your reasoning.

3

Characterizing the radiation field

In this chapter we give the basic definitions that characterize the radiation field. We start out with a short discussion of the two geometries that will be used frequently throughout these lecture notes.

3.1 Coordinate systems

Plane-parallel layers

Throughout these lectures we will almost exclusively concentrate on one-dimensional problems in either a geometry of plane-parallel layers or of spherical shells (see figure 3.1).

In a geometry of plane-parallel homogeneous layers we use a Cartesian coordinate system, such that $\mathbf{r} = (x, y, z)$ and z denotes the direction normal to the planar layers. In case of homogeneous layers *axial symmetry* will hold, such that all quantities relevant in describing the state of the medium depend on z only. They are constant in the x and y directions, therefore all gradients $\partial/\partial x$ and $\partial/\partial y$ will be zero. For a beam entering or exiting the medium at a polar angle θ between the normal direction and the beam direction, we will frequently use the variable $\mu = \cos\theta$.

The atmospheres of most stars and planets are well described using a plane parallel geometry. This may seem surprising. However, taking the sun as an example, one should realize that the thickness of the solar atmosphere (where the continuum and spectral lines in the visible part of the solar spectrum originate) is only some 500 km whereas the radius of the sun is about 700 000 km. For a sufficiently small patch of the solar atmosphere, its curvature is therefore negligible. For some stars the assumption of plane-parallel layers can not be justified. For example, if the star has a stellar wind (see chapter 17) the strongest lines may form over a radial extend of many (sometimes thousands of) stellar radii. In that case we must assume a geometry of spherical shells.

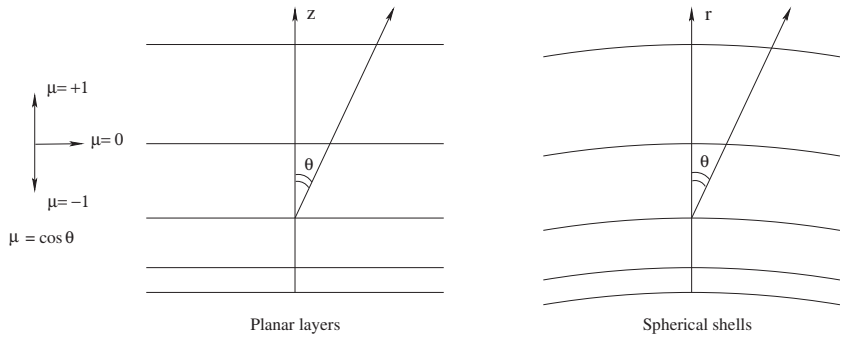


Figure 3.1: *Diagram showing the geometry of plane-parallel layers and of spherical shells.*

Spherical shells

In a spherical coordinate system the spatial point $\mathbf{r}$ is described using the coordinates (r, Θ, Φ). If we assume that the shells are homogeneous, *spherical symmetry* will hold. All quantities describing the medium will only depend on the radial coordinate r. At the position $\mathbf{r}$ we specify the direction of a beam using the polar and azimuthal angles (θ, ϕ). The angle θ is that between the radial direction and the beam direction, and we will frequently replace it by $\mu = \cos\theta$. As spherical symmetry implies axial symmetry, all azimuthal gradients $\partial/\partial\phi$ are zero.

Vector calculus in a geometry of planar layers or spherical shells

As a reminder we list some formulas from vector calculus used in these lecture notes to describe planar layers and spherical shells.

CARTESIAN COORDINATES $(\mathbf{x}, \mathbf{y}, \mathbf{z})$

A vector field	$\mathbf{A} = (A_x, A_y, A_z) = A_x\mathbf{x} + A_y\mathbf{y} + A_z\mathbf{z}$
Gradient	$\nabla f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right) = \frac{\partial f}{\partial x}\mathbf{x} + \frac{\partial f}{\partial y}\mathbf{y} + \frac{\partial f}{\partial z}\mathbf{z}$
Divergence	$\nabla \cdot \mathbf{A} = \frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z}$
Differential displacement	$d\mathbf{s} = dx\,\mathbf{x} + dy\,\mathbf{y} + dz\,\mathbf{z}$
Differential volume	$dV = dx\,dy\,dz$

SPHERICAL COORDINATES $(\mathbf{r}, \boldsymbol{\theta}, \boldsymbol{\phi})$

A vector field	$\mathbf{A} = (A_r, A_\theta, A_\phi) = A_r\mathbf{r} + A_\theta\boldsymbol{\theta} + A_\phi\boldsymbol{\phi}$
Gradient	$\nabla f = \left(\frac{\partial f}{\partial r}, \frac{\partial f}{\partial \theta}, \frac{\partial f}{\partial \phi}\right) = \frac{\partial f}{\partial r}\mathbf{r} + \frac{1}{r}\frac{\partial f}{\partial \theta}\boldsymbol{\theta} + \frac{1}{r\sin\theta}\frac{\partial f}{\partial \phi}\boldsymbol{\phi}$
Divergence	$\nabla \cdot \mathbf{A} = \frac{1}{r^2}\frac{\partial\left(r^2 A_r\right)}{\partial r} + \frac{1}{r\sin\theta}\frac{\partial(A_\theta \sin\theta)}{\partial \theta} + \frac{1}{r\sin\theta}\frac{\partial A_\phi}{\partial \phi}$
Differential displacement	$d\mathbf{s} = dr\,\mathbf{r} + rd\theta\,\boldsymbol{\theta} + r\sin\theta d\phi\,\boldsymbol{\phi}$
Differential volume	$dV = r^2 \sin\theta\, dr\, d\theta\, d\phi$

3.2 Specific intensity

The *specific intensity*, *surface brightness*, or *radiance* I_ν at position $\mathbf{r}$ and time t, traveling in direction $\mathbf{n}$, is defined such that the amount of energy transported by radiation of frequencies $(\nu, \nu + d\nu)$ across an element of area dS into a solid angle $d\omega$ in a time interval dt is

$$\begin{aligned} dE_\nu &= I_\nu(\mathbf{r}, \mathbf{n}, t)\ \mathbf{n} \cdot \mathbf{dS}\ d\omega d\nu dt \\ &= I_\nu(\mathbf{r}, \mathbf{n}, t)\ \cos\theta\ dS\ d\omega d\nu dt \end{aligned} \tag{3.1}$$

where θ is the angle between the direction of the beam and the normal to the surface, i.e. $\mathbf{n} \cdot \mathbf{dS} = \mathbf{n} \cdot \mathbf{s} dS = \cos\theta\ dS$ (see figure 3.2). The dimensions of I_ν are erg cm^{-2} sec^{-1} hz^{-1} sr^{-1} (recall that 1 erg = 10^{-7} J). From a macroscopic point of view, the specific intensity provides a complete description of the (unpolarized) radiation field.

Formally it would be better to refer to $I_\nu(\mathbf{r}, \mathbf{n}, t)$ as the *surface brightness*, however we will opt for the term *specific intensity*. By default we will drop an explicit reference to the position, time and direction for which the specific intensity is given and will suffice with I_ν. If the specific intensity is given as a function of wavelength, we will use I_λ. These quantities can be converted into each other if one realizes that

$$|I_\nu\, d\nu| = |I_\lambda\, d\lambda| \tag{3.2}$$

Using $\nu = c/\lambda$ yields: $I_\lambda = c/\lambda^2 I_\nu$. The dimensions of I_λ are erg cm^{-2} sec^{-1} cm^{-1} sr^{-1}.

The frequency integrated or *total specific intensity* is

$$I \equiv \int_0^\infty I_\nu\, d\nu = \int_0^\infty I_\lambda\, d\lambda \tag{3.3}$$

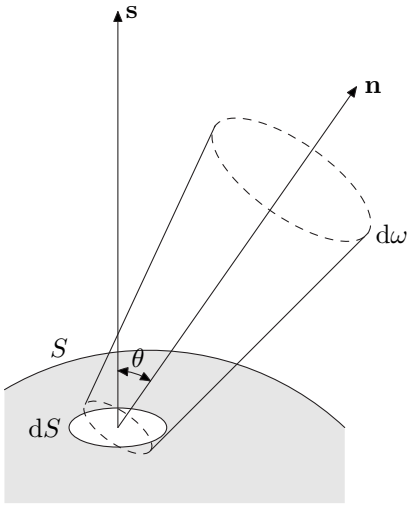


Figure 3.2: *Definition of the specific intensity as a beam of radiation into a solid angel $d\omega$, across a surface dS oriented in direction* $\mathbf{s}$*. The vector* $\mathbf{n}$ *is the direction of propagation of the beam which is at an angle θ with* $\mathbf{s}$*.*

The invariance of I_ν

The specific intensity is defined such that it is independent of distance if there are no sources or sinks of radiation along the direction of the beam. This implies that the value for the specific intensity at the source location can be obtained by measuring the amount of energy per unit time and per frequency interval that hits a detector (with known efficiency) per unit detector surface, when the solid angle subtended by the source is known. To be able to measure the specific intensity it is therefore required that the source is spatially resolved.

In other words: if we use our telescope to observe a fragment of an extended source (for instance a nebula, a galaxy, a planet, or the sun) somewhere in the sky, then the intensity that we derive from the amount of energy that reaches our detector, per unit frequency and time, is the same as that is emitted by the fragment in our direction.

We can understand this property by considering a beam of radiation that passes through a surface element dS at position $\mathbf{r}$, as well as through an element dS' at position $\mathbf{r}'$ (see figure 3.3). The amount of energy passing through both areas is

$$dE_\nu = I_\nu \cos\theta \, dS \, d\omega d\nu dt = I'_\nu \cos\theta' \, dS' \, d\omega' d\nu dt \tag{3.4}$$

where $d\omega$ is the solid angle subtended by dS' as seen from position $\mathbf{r}$, and $d\omega'$ is the solid angle subtended by dS as seen from $\mathbf{r}'$. Given that the distance between $\mathbf{r}$ and $\mathbf{r}'$ is equal to d, it follows from $d\omega = 4\pi \ \cos\theta' dS'/4\pi d^2$ and $d\omega' = 4\pi \ \cos\theta \, dS/4\pi d^2$ that $I_\nu = I'_\nu$.

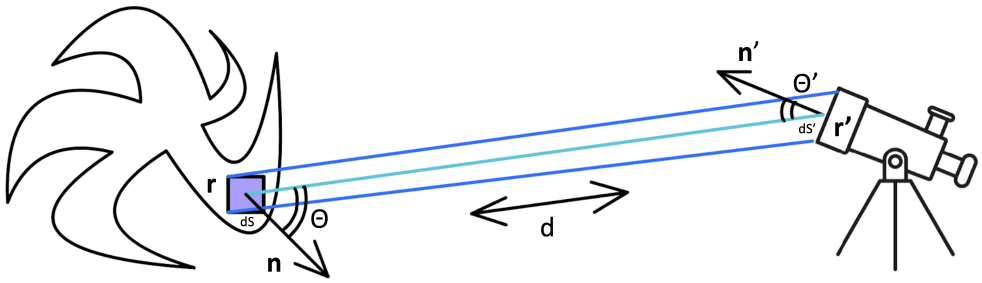


Figure 3.3: *Illustration of the proof of invariance of the specific intensity. The points* $\mathbf{r}$ *and* $\mathbf{r}'$ *are separated by a distance* d*. Area* dS *subtends a solid angle* $d\omega'$ *seen from position* $\mathbf{r}'$*, and the area* dS' *subtends* $d\omega$ *at* $\mathbf{r}$*.* $\mathbf{n}$ *and* $\mathbf{n}'$ *are unit vectors normal to* dS *and* dS'*.*

Lambertian emission and reflectance

A *Lambertian* surface emits (or reflects) specific intensity equally in all directions. Lambertian surfaces therefore are also called isotropically radiating surfaces. For stellar surfaces, Lambertian emission is a poor approximation (see 10.1) – though we will use it anyways for simple transfer problems. A Lambertian surface is often used to approximate the reflectivity of planetary bodies. If a planar Lambertian planetary surface illuminated from above reflects an amount of energy $dE_\nu(\theta = 0)$ from a surface element dS into a solid angle $d\omega$ in a frequency band $d\nu$ in a time interval dt in the direction normal to the surface, it will reflect an energy

$$dE_\nu(\theta) = \cos\theta \, dE_\nu(\theta = 0) = \mu \, dE_\nu(\mu = 1) \tag{3.5}$$

in the direction θ.

3.3 Mean intensity

The mean specific intensity or *mean intensity* averaged over all directions is

$$J_\nu(\mathbf{r}, t) = \frac{1}{4\pi} \oint I_\nu(\mathbf{r}, \mathbf{n}, t) \, d\omega \tag{3.6}$$

The mean intensity is the angular moment of order zero of the specific intensity and has dimensions erg cm^{-2} sec^{-1} hz^{-1}. In spherical coordinates $d\omega = \sin\theta d\theta d\phi = -d\mu d\phi$ (see figure 3.4). The total solid angle Ω therefore is

$$\Omega = \oint d\omega = \int_0^{2\pi} \int_0^{\pi} \sin\theta \, d\theta d\phi = 2\pi \int_{-1}^{+1} d\mu = 4\pi \tag{3.7}$$

This explains the normalisation factor 4π in eq. (3.6). Solid angle is expressed in the dimensionless unit steradian (sr). In the small angle limit, i.e. $\theta_\star \to 0$ or $\mu_\star = \cos\theta_\star \to 1$ of the

axially symmetric case (for instance, we observe a circular looking object in the sky), we find

$$\Omega_\star = 2\pi \int_0^{\theta_\star} \sin\theta \; d\theta = 2\pi \left(1 - \cos\theta_\star\right) \simeq 2\pi \left(1 - \left[1 - \theta_\star^2/2\right]\right) = \pi \, \theta_\star^2, \tag{3.8}$$

where $\theta_\star$ is in radians and the one-but-final equality follows from Taylor expansion of the $\cos\theta_\star$ term.

Notice that though the mean intensity is defined per steradian, through the factor $1/4\pi$ in eq. 3.6, its dimension does not convey this. The reasoning behind this is that J_ν is no longer a function of solid angle; as solid angle is a dimensionless unit explicit reference to it may be dropped. We may now rewrite eq. (3.6) as

$$\begin{aligned} J_\nu &= \frac{1}{4\pi} \int_0^{2\pi} \int_0^{\pi} I_\nu \sin\theta \; d\theta d\phi \\ &= \frac{1}{2} \int_{-1}^{+1} I_\nu \; d\mu \end{aligned} \tag{3.9}$$

where the latter equality assumes axial or spherical symmetry. The mean intensity is, for instance, used in the description of processes such as photoexcitation and photoionization, which only depend on the number of photons at some position at some time and do not depend on the direction of origin of these photons. For an isotropic radiation field $J_\nu = I_\nu$.

The frequency integrated or *total mean intensity* is

$$J \equiv \int_0^\infty J_\nu \; d\nu \tag{3.10}$$

Geometrical dilution

Consider a spherical star that emits an isotropic radiation field $I_\nu(\theta, \phi) = I_\nu$ from its surface at $R_\star$. We are interested in the mean intensity above the stellar surface, i.e. at $r > R_\star$. Figure 3.5 shows the directions from which the point r receives the stellar intensity I_ν. For the mean intensity in this point we find

$$J_\nu(r) = \frac{1}{2} \int_{\mu_\star}^1 I_\nu \, d\mu = \frac{1}{2}(1 - \mu_\star) \, I_\nu \equiv W(r) \, I_\nu \tag{3.11}$$

where $\mu_\star = \left[1 - (R_\star/r)^2\right]^{1/2}$ is the grazing angle that just hits the stellar rim (seen from r) and

$$W(r) = \frac{1}{2} \left\{ 1 - \left[1 - \left(\frac{R_*}{r} \right)^2 \right]^{1/2} \right\} \tag{3.12}$$

The factor W is called *geometrical dilution* and denotes the fraction of the total solid angle subtended by a star of radius $R_\star$ seen from a point in the sky at distance r. If $r = R_\star$ one finds

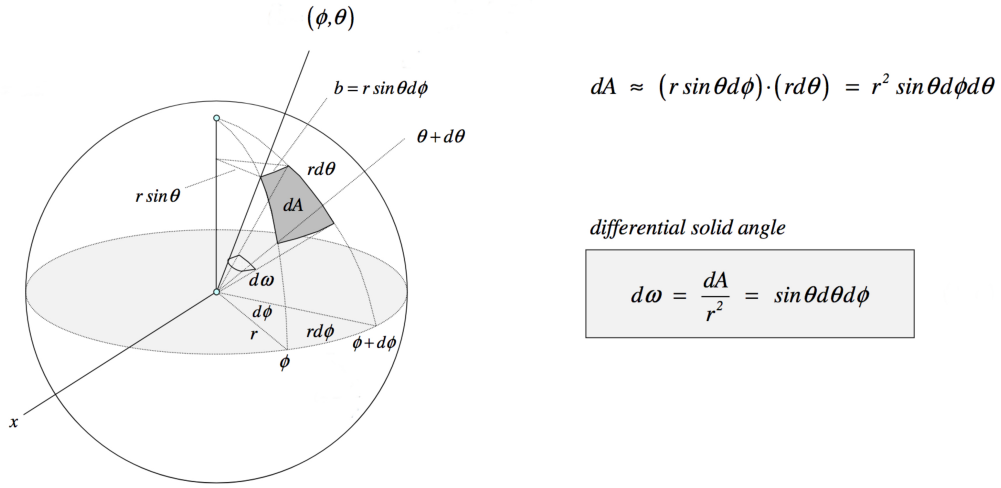


Figure 3.4: *The infinitesimal solid angle $d\omega$ expressed in spherical polar coordinates using colatitude or polar angle θ and longitude or azimuthal angle ϕ. The area A of a sphere of radius r in the interval $(\theta, \theta + d\theta)$ and $(\phi, \phi + d\phi)$ is $A = b \cdot a = r^2 \sin\theta d\theta d\phi = r^2 d\omega$, so that $d\omega = \sin\theta d\theta d\phi$. The total solid angle is 4π steradians.*

$W = 1/2$. This is easy to understand as an observer on the stellar surface sees half of the sky filled by the star. At very large distances, i.e. $r \gg R_\star$, one may approximate the dilution factor by

$$W(r) \simeq \frac{1}{4}\left(\frac{R_\star}{r}\right)^2 . \tag{3.13}$$

This also is easy to grasp: at large distances one sees the stellar disk, therefore $W(r) = \pi R_\star^2/4\pi r^2$.

Energy density of radiation

The monochromatic energy density of radiation is the amount of energy present in a region of space per unit volume in frequency range $(\nu, \nu + d\nu)$. In order to obtain the energy density u_ν of the radiation field (dimensions ergs cm^{-3} hz^{-1}) we determine the amount of radiation energy E_ν that is contained in an elementary volume V, such that $u_\nu = E_\nu/V$.

The amount of energy in a beam of solid angle $d\omega$ passing through a surface element dS of this volume is given by eq. (3.1). We are only interested in those photons in the beam that are actually inside V. Say the path length of these photons through V is l, then they will be inside of the elementary volume during a time $dt = l/c$. The part of the volume in which the photons are located is $dV = l\, dS \cos\theta$, where θ is the angle between the direction of the beam and the normal to the surface element dS. The energy in this part of the volume provided by the beam is therefore $dE_\nu = c^{-1}\, I_\nu(\mathbf{r}, \mathbf{n}, t)\, d\omega\, d\nu\, dV$. Integrating over all beams (to cover the entire

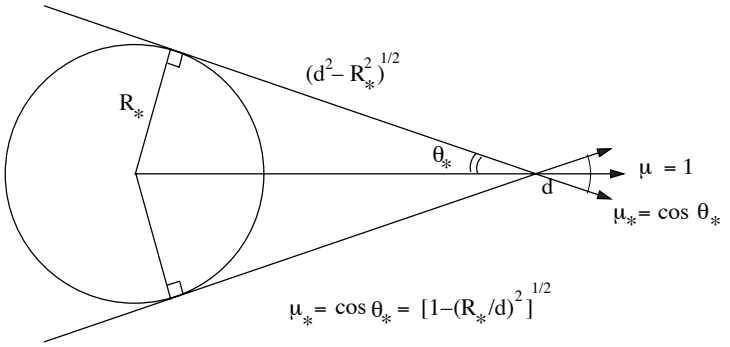


Figure 3.5: *Geometry showing the directions from which a point d receives the stellar intensity I_ν.*

volume) and the volume (to account for the size of the volume through the path length l), one finds for the total energy in V in the frequency band $d\nu$

$$E_\nu(\mathbf{r},t)\,d\nu = \frac{1}{c}\left[\int_V \left\{\oint I_\nu(\mathbf{r},\mathbf{n},t)\,d\omega\right\} dV\right] d\nu \tag{3.14}$$

As the volume V is chosen so small that it is elementary, i.e. I_ν is independent of position within V, we may evaluate the integrals independently. We then find for the *monochromatic energy density*

$$u_\nu(\mathbf{r},t) = \frac{E_\nu(\mathbf{r},t)}{V} = \frac{1}{c}\oint I_\nu(\mathbf{r},\mathbf{n},t)\,d\omega = \frac{4\pi}{c}J_\nu(\mathbf{r},t) \tag{3.15}$$

The frequency integrated *total energy density* is

$$u = \int_0^\infty u_\nu\,d\nu = \frac{4\pi}{c}\,J. \tag{3.16}$$

3.4 Flux

We define the flux of radiation $\boldsymbol{\mathcal{F}}_\nu$(**r**,t) as a vector quantity such that $\boldsymbol{\mathcal{F}}_\nu \cdot \mathbf{dS}$ gives the net rate of radiant energy flow across the arbitrarily oriented surface $\mathbf{dS} = \mathbf{s}\,dS$ per unit time and frequency interval.

The flux can be derived from the specific intensity passing through surface $\mathbf{dS}$ if we integrate over all solid angles. The energy that passes through the surface in frequency interval $d\nu$ in time dt can be written as

$$\boldsymbol{\mathcal{F}}_\nu(\mathbf{r},t)\cdot\mathbf{dS}\,d\nu\,dt = \oint dE_\nu = \oint I_\nu(\mathbf{r},\mathbf{n},t)\,\mathbf{n}\cdot\mathbf{dS}\,d\omega\,d\nu\,dt, \tag{3.17}$$

where the integration is over solid angle, consequently

$$\boldsymbol{\mathcal{F}}_\nu(\mathbf{r},t) = \oint I_\nu(\mathbf{r},\mathbf{n},t)\ \mathbf{n}\ d\omega, \tag{3.18}$$

that is

$$(\mathcal{F}_x,\mathcal{F}_y,\mathcal{F}_z) = \left(\oint I_\nu(\mathbf{r},\mathbf{n},t)\ \mathrm{n}_x\ d\omega,\ \oint I_\nu(\mathbf{r},\mathbf{n},t)\ \mathrm{n}_y\ d\omega,\ \oint I_\nu(\mathbf{r},\mathbf{n},t)\ \mathrm{n}_z\ d\omega\right). \tag{3.19}$$

The flux has dimensions erg cm^{-2} sec^{-1} hz^{-1}. One may therefore also think of the flux being the power per unit surface per unit frequency bandwidth (which a radio astronomer would likely find more appealing). In infrared and radio astronomy the flux is often given in units of jansky (symbol: Jy): 10^{-23} erg cm^{-2} sec^{-1} hz^{-1} $\equiv 1$ Jy.

In a plane-parallel medium only the flux in the z direction $\mathcal{F}_z \neq 0$. Symmetry arguments show that in the x and y directions the flux $\mathcal{F}_x = \mathcal{F}_y = 0$. As in this case only the z component of the flux is relevant one usually denotes this flux vector component as 'the' flux. It follows that

$$\begin{aligned}\mathcal{F}_\nu(z,t) &= \int_0^{2\pi}\int_0^{\pi} I_\nu \cos\theta \sin\theta\ d\theta d\phi \\ &= 2\pi\int_{-1}^{+1} I_\nu(z,\mu,t)\,\mu\ d\mu\end{aligned} \tag{3.20}$$

Again: $\mathcal{F}_\nu(z,t)$ is the net flow of radiant energy, per second per frequency interval, that passes through a surface of 1 cm^2 that at position z is oriented normal to the z-direction. It is a *net* flow of energy as the perspective factor $\mu = \cos\theta$ (measuring the effective surface) counts inward directed contributions ($-1 \leq \mu \leq 0$) negative. We may write

$$\begin{aligned}\mathcal{F}_\nu(z,t) &= 2\pi\int_0^{+1} I_\nu(z,\mu,t)\,\mu\ d\mu - 2\pi\int_0^{-1} I_\nu(z,\mu,t)\,\mu\ d\mu \\ &\equiv \mathcal{F}_\nu^+(z,t) - \mathcal{F}_\nu^-(z,t)\end{aligned} \tag{3.21}$$

where the outward directed flux $\mathcal{F}_\nu^+$ and the inward directed flux $\mathcal{F}_\nu^-$ are both positive. With the 'flux of a star' one most often intends to say the outward directed or *emergent flux* $\mathcal{F}_\nu^+$.

Two alternative definitions for the flux are also used. These are the *astrophysical flux* $F_\nu \equiv \pi^{-1}\mathcal{F}_\nu$ and the *Eddington flux*

$$H_\nu \equiv \frac{1}{4\pi}\mathcal{F}_\nu = \frac{1}{2}\int_{-1}^{+1} I_\nu\,\mu\ d\mu \tag{3.22}$$

which is of a similar form as eq. (3.9) for the mean intensity and is referred to as the angular moment of order one of the specific intensity.

One obtains the total energy passing through a unit surface, per second, by integrating the flux over frequency. This yields the *total flux*

$$\mathcal{F} \equiv \int_0^\infty \mathcal{F}_\nu\ d\nu \equiv \sigma\, T_{\rm eff}^4, \tag{3.23}$$

in erg cm^{-2} sec^{-1}. The last equality already introduces the *effective temperature* $T_{\rm eff}$ (see § 6.6). Integration of the flux over all of the stellar surface results in the *monochromatic luminosity*

$$L_\nu = 4\pi R_\star^2 \, \mathcal{F}_\nu(R_\star) \tag{3.24}$$

in erg sec^{-1} hz^{-1}, where $R_\star$ is the stellar radius. Further integration over all frequencies yields the *luminosity* of the star

$$L_\star = \int_0^\infty L_\nu \, d\nu = 4\pi R_\star^2 \, \mathcal{F}(R_\star) = 4\pi R_\star^2 \, \sigma \, T_{\rm eff}^4 \tag{3.25}$$

in erg sec^{-1}. Even though planets do not produce energy through thermo-nuclear reactions in their interiors, they too have a luminosity originating from gravitational potential energy left over from the planet's natal contraction and radioactive decay.

The r^{-2} dependence of the flux

The observational meaning of the flux can be understood in the following way: Consider a constant, isotropically radiating source, e.g. a spherical star with radius $R_\star$. When we place a concentric spherical surface around this source, with radius r, the total amount of radiation energy passing through the outer surface will be the same as that passing through the stellar surface, assuming there is no absorption or emission in the space around the source. Therefore

$$\mathcal{F}_\nu(R_\star) \, 4\pi R_\star^2 = \mathcal{F}_\nu(r) \, 4\pi r^2 \tag{3.26}$$

The flux therefore decreases with distance as $\mathcal{F}_\nu(r) \propto r^{-2}$. If d is the distance of the source to earth, then $\mathcal{F}_\nu(d)$ is the observed flux. For the sun, the observed flux is also referred to as the *solar irradiance* $\mathcal{R}_\nu$.

At first sight it may seem that this result is in contradiction to the invariance of the specific intensity along the line of sight. This is not so. Again consider a constant, isotropically radiating star with radius $R_\star$ and emerging intensity I_ν. At a point r the intensity is equal to I_ν for all beams that intersect the star, for all other beams it is zero. We then find for the flux in r (see figure 3.5)

$$\mathcal{F}_\nu(r) = 2\pi \int_{\mu_\star}^1 I_\nu \, \mu \, d\mu = \pi I_\nu (1 - \mu_\star^2) = \pi I_\nu \left(\frac{R_\star}{r} \right)^2 \tag{3.27}$$

So, we again find the flux to be proportional to r^{-2}. Note that the flux at the surface $\mathcal{F}_\nu(R_\star) = \pi I_\nu$, a result that can also be obtained directly from

$$\mathcal{F}_\nu(R_\star) = 2\pi \int_{\mu_\star=0}^1 I_\nu \, \mu \, d\mu \equiv \mathcal{F}_\nu^+(R_\star) = \pi I_\nu \tag{3.28}$$

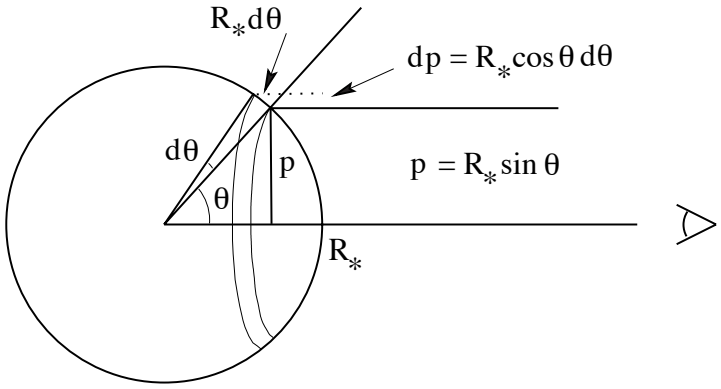


Figure 3.6: *Geometry of measurement of stellar flux. The projected annulus on the surface of the star has an area* $dS = 2\pi p dp = 2\pi R_\star \sin\theta\, R_\star \cos\theta\, d\theta = -2\pi R_\star^2\, \mu\, d\mu$*, normal to the line of sight. This area subtends a solid angle* $d\omega = dS/d^2$ *as seen by the observer at distance* d*.*

Reconstruction of the flux using a ray-by-ray description

We formulate the r^{-2} dependence of the flux in yet another way. Assume that the distance d between star and observer is very much larger than the stellar radius, i.e. $d \gg R_\star$, so that all rays of light from star to observer may be considered to be parallel. The energy received, per unit area normal to the line of sight, from a differential area on the star is

$$d\mathcal{F}_\nu(d) = I_\nu\, d\omega, \tag{3.29}$$

where $d\omega$ is the solid angle subtended by the area, seen from the position of the observer. This follows directly from the definition of the flux, eq. (3.18).

We subdivide the stellar surface in differential projected annuli $dS = 2\pi p\, dp = -2\pi R_\star^2\, \mu\, d\mu$, where $\mu = \cos\theta$ as usual (see figure 3.6). The projected distance p to the center of the stellar disk is referred to as the *impact parameter*. The solid angle of the projected annulus is $d\omega = dS/d^2 = -2\pi(R_\star/d)^2\, \mu\, d\mu$ as seen by the observer. The radiation emitted from this annulus in the direction of the observer emerged at angle θ; hence the appropriate value of the specific intensity is $I_\nu(\mu)$. Integrating over the disk, we find

$$\mathcal{F}_\nu(d) = \frac{2\pi}{d^2}\int_0^{R_\star} I_\nu(p)\, p\, dp = \left(\frac{R_\star}{d}\right)^2 2\pi \int_0^1 I_\nu(\mu)\, \mu\, d\mu = \left(\frac{R_\star}{d}\right)^2 \mathcal{F}_\nu^+(R_\star) \tag{3.30}$$

Once more we see that the flux is proportional to r^{-2}. It is because the solid angle subtended by the stellar disk decreases with the square of the distance. This way of reconstructing the flux is known as a *ray-by-ray* solution.

Angular diameter

The *angular diameter* of an object (e.g., a star, the moon) at distance d is

$$\alpha_\star = \frac{2R_\star}{d} \tag{3.31}$$

such that the observed flux can also be written as $\mathcal{F}_\nu(d) = (\alpha_\star^2/4)\,\mathcal{F}_\nu(R_\star)$. For spatially resolved objects, we may directly constrain the stellar radius if the distance is known, $R_\star = \alpha_\star\, d/2$, as well as the flux at the objects's surface, $\mathcal{F}_\nu(R_\star)$. If we know the latter quantity from (other) measured quantities we may determine the angular diameter of the object, even if we can't measure $\alpha_\star$ directly. This too leads to an estimate of the stellar radiuis.

3.5 Photon momentum (a.k.a. radiation pressure)

Consider an amount of momentum p_ν transported by the radiation field through an area dS, per unit time per frequency interval.

A photon of frequency ν has a momentum $h\nu/c$ such that the amount of momentum in a beam with energy dE_ν is dE_ν/c. If we consider the component of the momentum in the direction normal to that of surface dS, then the contribution of the beam is

$$dp_\nu\, dS d\nu dt = \frac{1}{c} dE_\nu \cos\theta \tag{3.32}$$

where again θ is the angle between the direction of the beam $\mathbf{n}$ and the direction $\mathbf{s}$ normal to the surface. If we express the energy in terms of the specific intensity, using eq. (3.1), and if we integrate over all directions we find

$$\begin{aligned} p_\nu &= \frac{1}{c} \oint I_\nu(\mathbf{r}, \mathbf{n}, t) \mathbf{nn}\, d\omega \\ &= \frac{1}{c} \oint I_\nu \cos^2\theta\, d\omega \end{aligned} \tag{3.33}$$

The dimensions of p_ν are ergs cm^{-3} hz^{-1}. For clarity: the one "$\cos\theta$" term determines the effective size of area dS, oriented in the direction $\mathbf{s}$, for radiation propagating in direction $\mathbf{n}$; the other "$\cos\theta$" determines the momentum component in the direction $\mathbf{s}$.

If axial symmetry is valid

$$p_\nu = \frac{2\pi}{c} \int_{-1}^{+1} I_\nu(z, \mu, t)\, \mu^2\, d\mu = \frac{4\pi}{c} K_\nu \tag{3.34}$$

where

$$K_\nu \equiv \frac{1}{2} \int_{-1}^{+1} I_\nu\, \mu^2\, d\mu \tag{3.35}$$

is the angular moment of order two of the specific intensity.

The quantity p_ν is called the *radiation pressure*, which is highly confusing as it is *not* the mechanical force per unit surface that would be exerted by photons on an absorbing screen. We may derive this mechanical force from the above equation though, realizing that radiation only exerts a force when there is a gradient in the radiation pressure – in analogy to the force associated with a gas pressure gradient. We will return to this when we discuss hydrostatic equilibrium (see § 9.3). Radiation pressure is a scalar when the radiation field is isotropic.

The total radiation pressure is

$$p_{\rm R} = \int_0^\infty p_\nu \, d\nu \tag{3.36}$$

where we have introduced the subscript R to avoid confusion with the gas pressure (see e.g. eq. 6.28).

3.6 Eddington factors

The moment equations (3.9), (3.22) and (3.35) of the radiation field are used to define the so-called *Eddington factors*

$$f_\nu(\mathbf{r},t) \equiv \frac{K_\nu(\mathbf{r},t)}{J_\nu(\mathbf{r},t)} \tag{3.37}$$

and

$$g_\nu(\mathbf{r},t) \equiv \frac{H_\nu(\mathbf{r},t)}{J_\nu(\mathbf{r},t)}. \tag{3.38}$$

The Eddington factors have no dimension and vary within only a very limited range: f_ν typically between 1/3 and 1; and g_ν typically between 0 and 1. In essence both factors are a measure of the amount of anisotropicity that is contained in the radiation field. Is the radiation field completely isotropic then $f_\nu = 1/3$ and $g_\nu = 0$. Is the radiation field sharply peaked then both factors tend toward unity. In § 4.5 we will see that f_ν can be used to seemingly reduce the number of independent variables in a transfer problem. They also provide a means to close a system of moment equations derived from the equation of transfer. Later on, this will allow us to obtain a transfer equation for J_ν rather than for I_ν, a result that will be pivotal for obtaining the temperature structure of atmospheres in chapter 10. The factor g_ν is often used in formulating boundary conditions to the equation of transfer (see § 11.2).

Exercise 3.1

Assume that the sun emits an isotropic radiation field of which the specific intensity is I_ν. Express the solar specific intensity, mean intensity, flux, and irradiation as measured at the position of earth in I_ν. (No need to prove these relations; simply look them up in the lecture notes).

Exercise 3.2

From Earth, the mean distance to the moon is 385 000 km. For this distance, the lunar angular diameter is 31.02' (or 31'1").

a) Compute the radius of the moon.

b) Compute the geometrical dilution of the moon, as seen from Earth (pretend you are at the center of Earth for the remainder of this exercise).

c) Compute the solid angle $\Delta\Omega$ subtended by the moon. How many moons (at the mean lunar distance) are needed to fill the entire sky? (It is okay to use scissors to cut and paste lunar disks to fill holes in between full lunar disks).

Why is the Moon exactly the same apparent size from Earth as the Sun? Surely this cannot be just coincidence; the odds against such a perfect match are enormous.

Malcolm Smith, Hertfordshire, England

Published: Sunday, October 1, 2000

October 2000
Believe it or not, it actually is just a coincidence — and a happy one at that. The Moon and Sun have virtually the same angular size in our sky because the Sun is about 400 times wider than the Moon, but it's also about 400 times farther away. This allows us to see spectacular coronal displays and prominences during total solar eclipses. Interestingly, this hasn't always been the case. Tidal interactions cause the Moon to spiral about one inch per year away from Earth. In the distant past, the Moon was close enough to Earth so that it could block the Sun's entire disk and then some. Our prehuman ancestors would not have witnessed the beautiful coronal displays that we now enjoy. And about 50 million years from now, the Moon will be far enough away so that our descendants will only see annular eclipses.

Exercise 3.3

The *Total Irradiance Monitor* (TIM) on board the *Solar Radiation and Climate Experiment* (SORCE) measured the total solar irradiance above the Earth atmosphere to be $\mathcal{R} = 1.3608 \pm 0.0005 \times 10^6$ erg sec^{-1} or $1360.8 \pm 0.5W$ m^{-2} (Kopp & Lean, 2011). We measure the angular diameter of the sun to be 31.97' (or 31'.58"). Compute the effective temperature of the sun.

Exercise 3.4

At some point in the calculation of the observed stellar flux under the assumption of plane-parallel layers one needs to introduce the fact that a star is a sphere. Where is this done?

Exercise 3.5

Consider an infinately extended flat plate emitting from its surface an isotropic radiation field $I_\nu(\theta, \phi) = I_\nu$. Show that the mean intensity above the surface is given by $J_\nu = 1/2\ I_\nu$. Explain this is a simple way.

Exercise 3.6

This could be a nice exam question. We take a picture of the lunar disk, which is at distance d and emits an isotropic radiation field with specific intensity I_ν. Our camera has a spatial resolving power $d\omega \ll \pi R_{\rm m}^2/d^2$ – defined by the pixel size of the ccd – and we find that we need to open the shutter a time t to assure that we have the desired exposure.

a) If we would now place the moon at twice the original distance, how long should we open the shutter to gain an identical exposure?

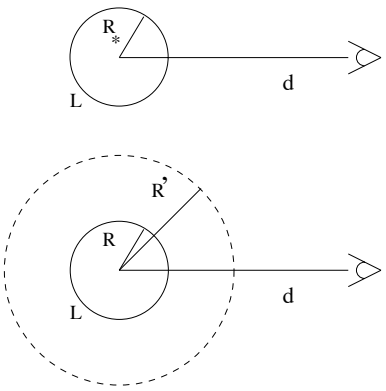
We now place the moon at such a distance that the lunar disk subtends a solid angle equal to the spatial resolving power of our camera, i.e. $d\omega = \pi R_{\rm m}^2/d^2$. The camera is pointed such that the moon falls in a single pixel.

b) How long should we open the shutter to get an identical exposure?

c) If now place the moon at twice the distance (than in b), implying we can no longer resolve the lunar disk. All photons that reach the ccd hit the same pixel. What should be the shutter time to get an identical exposure?

Exercise 3.7

L is a spatially resolved spherical lamp with radius R that is positioned at distance d from an observer and that radiates an isotropic radiation field with specific intensity I_ν.



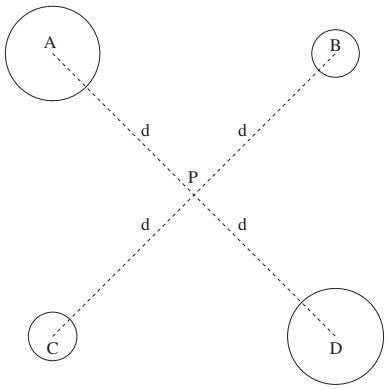
a) Give the specific intensity $I_\nu(d)$ and flux $\mathcal{F}_\nu(d)$ as measured by the observer.

It starts to fog. The light from the lamp is purely scattered with the result that it appears to originate from a spherical surface of radius $R\prime$.

b) For this new situation, give the specific intensity $I_\nu(d)$ and flux $\mathcal{F}_\nu(d)$ as measured by the observer.

Exercise 3.8

Consider a cluster of four stars; each star radiates an isotropic radiation field of which the specific intensity is I_ν. The four stars and the observer are in the same plane. The observer is in the center P of the cluster. The distance from P to each of the four stars is d. The radius of the stars B and C is R; that of the stars A and D is $2R$.

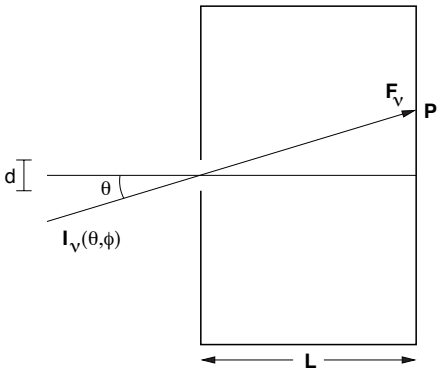


a) Give the mean intensity at P.

b) Give the flux at P.

Exercise 3.9

This could be a nice exam question. A "pinhole" camera consists of a very small circular hole of diameter d, at distance L from the "film-plane" (see figure). The pinhole is so small that the specific intensity that falls through the pinhole on a square centimeter of the film-plane may be assumed to be constant.



a) Derive an expression for the solid angle $\Delta\Omega$ that is extended by the pinhole as seen from the position P on the image plane.

The energy received per unit area normal to the line of sight from solid angle $\Delta\Omega$ is given by $d\mathcal{F}_\nu = I_\nu(\theta, \phi)\Delta\Omega$.

b) Derive an expression for the flux $\mathcal{F}_\nu$ at the film plane.

Exercise 3.10

Show that for isotropic radiation $p_\nu = u_\nu/3$. What is the meaning of the factor 1/3?

Exercise 3.11

a) Deep in the atmosphere the radiation field will be almost isotropic, i.e. $I_\nu(\theta, \phi) = I_\nu$. Show that in this case the Eddington factor $f_\nu = 1/3$.

b) Far above the stellar surface the radiation field will be strongly peaked in the radial direction, i.e. $I_\nu(\theta, \phi) = I_\nu(\mu, \phi) = I_\nu \delta(\mu - 1)$. Show that in this case the Eddington factor $f_\nu = 1$.

We may conclude that the Eddington factor will not vary much throughout the atmosphere (only by about a factor of three).

4

The equation of transfer

In this chapter we formulate the equation that describes the transfer of radiation through a medium, and we introduce the macroscopic quantities that play a role in this equation. The transfer equation has a formal solution which reflects that the specific intensity at each point in the medium can be determined if the source function S_ν and the optical depth τ_ν are known throughout the medium, or – equivalently – the extinction coefficient χ_ν and the emission coefficient η_ν. We discuss analytical solutions to simple transfer problems.

4.1 Absorption, emission, and scattering processes

Photons traveling through a material medium will experience interactions with the particles that are present: they can be absorbed or scattered. The medium itself may also emit radiation. Here, we give a macroscopic description of these processes. In principle, the nature of the medium is irrelevant. It may be a neutral or (partly) ionized gas, a molecular gas, or a medium of solid state particles.

We will make a distinction between "real" absorption and emission processes on the one hand, and scattering processes on the other hand. In a "real" absorption process (for the sake of convenience we will simply refer to this as absorption) energy is removed from the radiation field, and added to the local thermal energy of the medium. In this process photons are truely destroyed and we say that the photon has been thermalized. The reverse process is thermal emission, which adds energy from the local thermal pool into the radiation field. The photons that are created in this way are referred to as *thermal* photons. An example of a 'real' absorption process is the photo excitation of an electron in an atom or ion, followed by a downward transition in which the energy that is released is added to a free electron that is interacting with the atom or ion. An example of a thermal emission is the reverse process, i.e. a collisional excitation involving a free electron followed by the spontaneous emission of a photon.

In a scattering process the photon essentially only changes direction. So, although the photon

is absorbed and emitted, it keeps, as it were, its identity. One distinguishes between *pure* and *non pure* scatterings. In a pure scattering the frequency of the photon changes only slightly. Examples of pure scattering processes are the interaction of photons in a resonance transition (for instance Lyα or C IV $\lambda\lambda$1548,1551) or the elastic scattering of photons by free electrons (Thomson scattering) or by an atom or molecule (Rayleigh scattering). In a scattering process that is not pure the frequency of the photon is changed significantly. An example of such a process is the in-elastic scattering of high-energy photons by free electrons (Compton scattering).

When "real" absorption and emission processes dominate over scattering processes there will be a strong coupling between the local thermodynamic properties of the gas and the properties of the radiation that is generated by it. Concerning scattering processes the situation is different. Here the nature of the emitted radiation will predominantly be determined by the properties of the infalling radiation field. If scatterings dominate, the photons that are present in the local medium typically originate from elsewhere and simply propagate through the local medium without actually being coupled to it in any way. Depending on the origin of the photons the radiation field may be thermal or non-thermal.

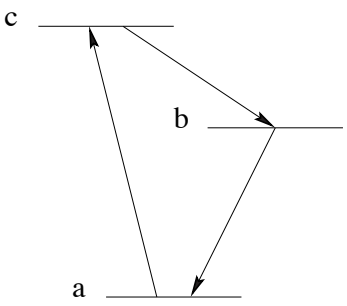


Figure 4.1: *Photon conversion*

A description in terms of absorption/emission and of scattering has its conceptual limitations. Take for instance an atom that has three bound states, a, b, and c, in order of increasing energy (see figure 4.1). Say that an absorption of a photon brings the atom from state a to c, and that this is followed by the spontaneous emission of a photon bringing the atom in state b, after which the atom returns into state a after again emitting a photon. One can not speak of a "real" absorption/emission as the thermal energy content of the gas did neither increase nor decrease. Also, it is clear that one can not speak of a scattering as the initial photon lost its identity.

Processes in a beam of light

The transfer of radiation along a ray of light – in terms of "real" absorption and emission and pure scattering – is schematically given in figure 4.2. There are two processes that add photons of given frequency to the ray, i.e. "real" emissions and scattering of photons that come in from different directions into the direction of the ray. There are also two processes that remove photons from the ray, i.e. "real" absorptions and scattering of photons out of the direction of the beam.

We will now formulate a quantitative description of these processes.

— — —

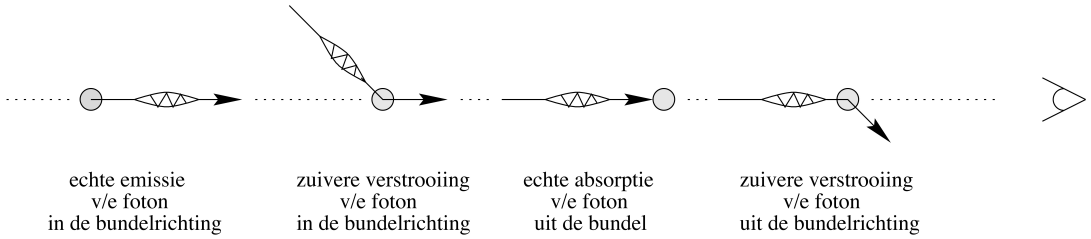


Figure 4.2: *Schematic representation of the transfer of radiation along a ray of light.*

Let us consider a gas and a radiation field that is described by a specific intensity $I_\nu(s)$ in a given direction. The path-length along the ray is described by s. Let us consider a flat cylinder, of length ds and cross-section dO, perpendicular to s. We describe changes in $I_\nu(s)$ due to absorption and emission processes between s and $s+ds$ as follows

$$\begin{aligned} I_\nu(s+ds)\, dO\, d\omega\, d\nu\, dt &= I_\nu(s)\, dO\, d\omega\, d\nu\, dt + dI_\nu(s)\, dO\, d\omega\, d\nu\, dt \\ \text{or} & \\ I_\nu(s+ds) &= I_\nu(s) + dI_\nu(s) \end{aligned} \tag{4.1}$$

Extinction coefficient

If only extinction processes occur along the path-length ds, either as a result of absorption or because of scattering out of the direction of the ray, then the decrease in the specific intensity is proportional to the incident specific intensity, to the path-length ds, and to the properties and number of absorbing/scattering particles. The constant of proportionality is called the extinction coefficient and can be defined in three ways (see eq. 4.5 for an overview).

We will often use the *linear extinction coefficient* χ_ν, which has dimension cm^{-1}, such that

$$dI_\nu(s) = -\chi_\nu(s)\, I_\nu(s)\, ds \tag{4.2}$$

This coefficient is defined such that an element of material, of cross-section dO and length ds, removes from a beam with specific intensity $I_\nu(\mathbf{r},\mathbf{n},t)$, incident normal to dS and propagating into a solid angle $d\omega$, an amount of energy

$$dE_\nu \equiv \chi_\nu(\mathbf{r},\mathbf{n},t)\, I_\nu(\mathbf{r},\mathbf{n},t)\, dO\, ds\, d\omega\, d\nu\, dt \tag{4.3}$$

within a frequency band $d\nu$ in a time dt. The linear extinction coefficient is sometimes also called the *volume extinction coefficient* as it refers to the total extinction surface in cm^2, per cm^{-3} ([cm^2 cm^{-3}] = [cm^{-1}]). The value of the extinction coefficient is determined by the properties/state of the medium and is not dependent on the radiation field. However, the latter can have an indirect effect on the properties/state of this medium.

The linear extinction coefficient can be split up in a contribution due to absorption, κ_ν, and due to scattering, σ_ν

$$\chi_\nu(s) = \kappa_\nu(s) + \sigma_\nu(s) \tag{4.4}$$

Here it is implicitly assumed that these processes are independent of each other and that they are additive (see eq. 4.9).

Alternative definitions of the extinction coefficient are

$$\chi_\nu = \chi'_\nu \, \rho = \alpha_\nu \, n \tag{4.5}$$

Here χ'_ν is the *mass extinction coefficient* ([cm^2 gr^{-1}]) and ρ the density of the material medium ([gr cm^{-3}]). α_ν is the extinction *cross section* per particle ([cm^2]) and n the number density of particles ([cm^{-3}]) that cause the extinction.

Emission coefficient

If only emission processes occur along the path-length ds, either as a result of thermal emission or because of scattering into the direction of the ray, then the increase in the specific intensity is given by

$$dI_\nu(s) = \eta_\nu(s) \, ds \tag{4.6}$$

The constant of proportionality is called the *volume emission coefficient* and is defined such that the amount of energy released from an element of material of cross-section dO and length ds, into a solid angle $d\omega$, within a frequency band $d\nu$, in direction $\mathbf{n}$ in a time interval dt, is

$$dE_\nu \equiv \eta_\nu(\mathbf{r}, \mathbf{n}, t) \; dO \, ds \, d\omega \, d\nu \, dt \tag{4.7}$$

The emission coefficient has dimensions [erg cm^{-3} sec^{-1} hz^{-1} sr^{-1}]. η_ν always depends on the properties of the medium, and, in case of scattering from other directions into the ray of light, also on the radiation field. For completeness we mention that in studies of gaseous nebulae it is custom to use the letter j (so j_ν) to denote the emission coefficient (for instance in Osterbrock & Ferland).

An alternative definition of the emission coefficient is

$$\eta_\nu = \eta\prime_\nu \rho \tag{4.8}$$

where $\eta\prime_\nu$ is the *mass emission coefficient* ([erg gr^{-1} sec^{-1} hz^{-1} sr^{-1}]).

— — —

If different extinction/emission processes occur at once they may be added

$$\chi_\nu = \sum_i \chi_{\nu,i} \qquad\qquad \eta_\nu = \sum_i \eta_{\nu,i} \tag{4.9}$$

The subscript i labels all different processes contributing at frequency ν.

Angular phase function for coherent scattering

For the emission coefficient of a coherent scattering process, i.e. where the frequency of the incoming photon ν' equals the frequency of the outgoing photon ν, we may write

$$\eta_\nu^{\rm sc}(\mathbf{n}) = \sigma_\nu \frac{1}{4\pi} \oint g(\mathbf{n}', \mathbf{n})\, I_\nu(\mathbf{n}')\, d\omega' \tag{4.10}$$

where σ_ν is the total scattering coefficient and g is the probability that a photon will be scattered from direction $\mathbf{n}'$ in solid angle $d\omega'$ into solid angle $d\omega$ in direction $\mathbf{n}$. This angular phase function is normalised to unity, i.e.

$$\frac{1}{4\pi} \oint g(\mathbf{n}', \mathbf{n})\, d\omega' = 1 \tag{4.11}$$

For isotropic scattering $g(\mathbf{n}', \mathbf{n}) \equiv 1$, and the emission coefficient for scattering reduces to $\eta_\nu^{\rm sc} = \sigma_\nu J_\nu$ (see Eq. 3.6). If in the interaction both the direction and frequency of the photon change, one introduces a redistribution function $R(\nu', \mathbf{n}'; \nu, \mathbf{n})$ to describe the process.

Extinction and emission: the equation of transfer

If both extinction and emission occur along the path-length ds then we may write

$$dI_\nu(s) = [\, \eta_\nu(s) - \chi_\nu(s)\, I_\nu(s)\,]\, ds \tag{4.12}$$

or

$$\frac{{\rm d}I_\nu}{{\rm d}s} = \eta_\nu - \chi_\nu\, I_\nu, \tag{4.13}$$

where in the latter equation we have dropped explicit reference to the spatial coordinate s. This is a simple form of the *equation of transfer*.

4.2 General form of the equation of transfer

Now that we know the quantities that describe the macroscopic interaction between radiation and matter, it is relatively simple to give a heuristic derivation of the general form of the equation of transfer. Let us again consider a ray of light of specific intensity I_ν, that is transported in a frequency interval $(\nu, \nu + d\nu)$, passing in a time dt through a volume element of length ds (such that $dt = ds/c$) and cross-section dO oriented normal to the ray in direction $\mathbf{n}$ into solid angle $d\omega$ (eq. 3.1). The difference between the amount of energy that emerges (at position $\mathbf{r} + \Delta\mathbf{r}$ and at time $dt + \Delta t$) and that incident (at $\mathbf{r}$ and t) must equal the amount created by emission from the material in the volume minus the amount absorbed. Therefore, using the

definition of specific intensity (eq. 3.1) and of the extinction and emission coefficient it must hold that in order to conserve the total photon energy

$$\begin{aligned}&[I_\nu(\mathbf{r}+\Delta\mathbf{r},\mathbf{n},t+\Delta t)-I_\nu(\mathbf{r},\mathbf{n},t)]\;dO\,d\nu\,dt\,d\omega=\\&[\eta_\nu(\mathbf{r},\mathbf{n},t)-\chi_\nu(\mathbf{r},\mathbf{n},t)\,I_\nu(\mathbf{r},\mathbf{n},t)]\;ds\,dO\,d\nu\,dt\,d\omega\end{aligned}\tag{4.14}$$

The difference in specific intensity at the left hand side of this equation can be written as

$$\begin{aligned}I_\nu(\mathbf{r}+\Delta\mathbf{r},\mathbf{n},t+\Delta t)-I_\nu(\mathbf{r},\mathbf{n},t) &= \frac{\partial I_\nu}{\partial t}dt+\frac{\partial I_\nu}{\partial s}ds\\ &= \left[\frac{1}{c}\frac{\partial I_\nu}{\partial t}+\frac{\partial I_\nu}{\partial s}\right]ds\end{aligned}\tag{4.15}$$

The general form of the equation of transfer holds for all geometries. To adopt a specific coordinate system one should express the derivatives along the ray of light in terms of the orthogonals in the chosen coordinate system. If ds is an infinitesimal displacement along the direction of the ray then

$$\frac{\partial}{\partial s}=\frac{\partial x}{\partial s}\frac{\partial}{\partial x}+\frac{\partial y}{\partial s}\frac{\partial}{\partial y}+\frac{\partial z}{\partial s}\frac{\partial}{\partial z}=n_x\frac{\partial}{\partial x}+n_y\frac{\partial}{\partial y}+n_z\frac{\partial}{\partial z}=\mathbf{n}\cdot\nabla,\tag{4.16}$$

for a Cartesian coordinate system, where (n_x,n_y,n_z) are the components of the unit vector $\mathbf{n}$.

For the general form of the equation of transfer we then find

$$\left(\frac{1}{c}\frac{\partial}{\partial t}+\mathbf{n}\cdot\nabla\right)I(\mathbf{r},\mathbf{n},\nu,t)=\eta(\mathbf{r},\mathbf{n},\nu,t)-\chi(\mathbf{r},\mathbf{n},\nu,t)\,I(\mathbf{r},\mathbf{n},\nu,t)\tag{4.17}$$

In these lectures we will assume by default that the medium is time-independent, in which case the partial derivative $\partial/\partial t=0$. Furthermore, we will focus on solving transfer equations in three special types of geometries. These are: *(i)* along a line, i.e. a pencil beam; *(ii)* a medium of homogeneous plane-parallel layers; or *(iii)* a medium of homogeneous spherical shells.

Along a pencil beam

The equation of transfer along a pencil beam, where s is the coordinate along the beam, is given by eq. (4.13).

Planar layers

If the medium consists of homogeneous plane-parallel layers, such that the properties of this medium are only a function of the z direction, then $n_z=dz/ds=\cos\theta=\mu$ (see figure 3.1).

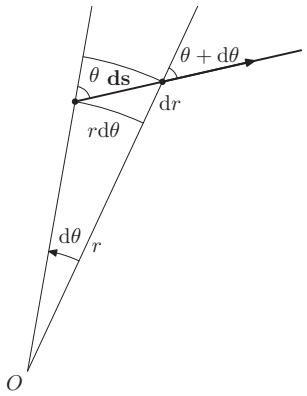


Figure 4.3: *Geometric relation among the variable* $\mathbf{ds}$ *and* dr *and* $d\theta$ *used in the derivation of the transfer equation in a spherically symmetric medium.*

The equation of transfer reduces to

$$\mu \frac{\mathrm{d}I_\nu(z,\mu)}{\mathrm{d}z} = \eta_\nu(z,\mu) - \chi_\nu(z,\mu)\, I_\nu(z,\mu) \tag{4.18}$$

Note that in the case $\mu = 1$, i.e. a beam along the z direction, the equation of transfer is equal to eq. (4.13).

Spherical shells

If the medium consists of homogeneous spherical shells, and θ and ϕ specify the direction of the beam at position r, such that θ is the angle of the beam relative to the local outward normal direction (which is the radial direction) and ϕ is the azimuthal angle, then the specific intensity will be independent of ϕ, i.e. $\partial/\partial\phi = 0$.

An infinitesimal displacement along the beam direction is given by $\mathbf{ds} = dr\ \hat{\mathbf{r}} + rd\theta\ \hat{\boldsymbol{\theta}}$. The geometry implies (see figure 4.3) that $dr = \cos\theta ds$ and $rd\theta = -\sin\theta ds$. Note that always $d\theta \leq 0$ for *any* ds; this explains the minus sign. We find

$$\frac{\partial}{\partial s} = \frac{\partial r}{\partial s}\frac{\partial}{\partial r} + \frac{\partial \theta}{\partial s}\frac{\partial}{\partial \theta} = \cos\theta \frac{\partial}{\partial r} - \frac{1}{r}\sin\theta \frac{\partial}{\partial \theta} = \mu \frac{\partial}{\partial r} + \frac{1}{r}\left(1-\mu^2\right)\frac{\partial}{\partial \mu} \tag{4.19}$$

where $\mu = \cos\theta$, such that

$$\frac{\partial}{\partial \theta} = \frac{\partial \mu}{\partial \theta}\frac{\partial}{\partial \mu} = -\sin\theta \frac{\partial}{\partial \mu} \tag{4.20}$$

Using these results the transfer equation for a spherically symmetric medium is

$$\left[\mu \frac{\partial}{\partial r} + \frac{1}{r}\left(1-\mu^2\right)\frac{\partial}{\partial \mu}\right] I_\nu(r,\mu) = \eta_\nu(r,\mu) - \chi_\nu(r,\mu)\, I_\nu(r,\mu) \tag{4.21}$$

4.3 Optical depth and the source function

Optical depth

Let us introduce an elementary *optical depth interval* $d\tau_\nu$ along a path-length ds, measured along the direction of the ray.

$$d\tau_\nu \equiv \chi_\nu(\mathbf{r}, \mathbf{n})\, ds \tag{4.22}$$

This defines the *optical depth* τ_ν which gives the integrated extinction of the material along the line of sight. For a geometrical thickness D the optical depth is

$$\tau_\nu(D) = \int_0^D \chi_\nu(s)\, ds \tag{4.23}$$

The optical depth is a dimensionless number. The physical meaning of τ_ν is simple. In the absence of emission in the medium the equation of transfer along the beam is given by (see eq. 4.2 and/or 4.13)

$$\frac{\mathrm{d}I_\nu}{\mathrm{d}\tau_\nu} = -I_\nu \tag{4.24}$$

This yields

$$I_\nu(D) = I_\nu(0)\, e^{-\tau_\nu(D)} \tag{4.25}$$

and shows that $\tau_\nu(D)$ is the exponential decline parameter that determines what remains of a beam that has passed through a layer of thickness D in which extinction processes occur. How far can photons penetrate in this layer? The chance that an incident photon travels an optical depth $\tau_\nu(s)$ in the layer (for $s < D$) is $p(\tau_\nu) = \exp(-\tau_\nu)$, where we have dropped explicit reference to s. The chance that after traveling this far it is absorbed within the interval $(\tau_\nu, \tau_\nu + d\tau_\nu)$ is $1 - \exp(-d\tau_\nu) \simeq d\tau_\nu$ for $d\tau_\nu \ll 1$. So, the probability that a photon will be scattered or absorbed between τ_ν and $\tau_\nu + d\tau_\nu$ is $\exp(-\tau_\nu)\, d\tau_\nu$. As the average of a quantity x that has a probability distribution $p(x)$ is given by

$$\langle x \rangle = \frac{\int_0^\infty x\, p(x)\, dx}{\int_0^\infty p(x)\, dx} \tag{4.26}$$

the *mean optical photon path* of the photon must be

$$\langle \tau_\nu \rangle = \frac{\int_0^\infty \tau_\nu\, e^{-\tau_\nu}\, d\tau_\nu}{\int_0^\infty e^{-\tau_\nu}\, d\tau_\nu} = 1 \tag{4.27}$$

Photons therefore typically travel one optical depth unit before interacting with the medium.

Mean free path of the photon

This result immediately shows what the mean geometrical path ([cm]) of a photon in a homogeneous medium must be, i.e. the path length it can travel before it is absorbed or scattered. This is

$$\ell_\nu = \frac{\langle \tau_\nu \rangle}{\chi_\nu} = \frac{1}{\chi_\nu} = \frac{1}{\chi'_\nu \rho} = \frac{1}{\alpha_\nu \, n} \tag{4.28}$$

and is formally refered to as the photon *mean-free-path*.

Optically thick and optically thin

One of the first questions one should ask oneself when trying to understand the transfer of radiation through a medium is whether this medium is *optically thin* or *optically thick* for radiation of given frequency. Optically thin implies that almost all incident photons, as well as almost all photons emitted in the medium, can propagate undisturbed by the medium and escape from it. An optically thin medium therefore describes the limiting case $\tau_\nu \ll 1$. In an optically thick medium $\tau_\nu \gg 1$. This (opposite) limit therefore describes a medium that is almost opaque for photons. Almost all photons in a beam that is incident to a medium that is optically thick will be absorbed or scattered inside the medium. Of all photons that are emitted by the medium only those will contribute to the emerging specific intensity that are emitted at a position and into a direction for which the optical depth to the edge of the medium is $\tau_\nu \sim 1$.

At which optical depth do half of the emitted photons escape? This is from $\tau_\nu = -\ln 0.5 = 0.693 \simeq 2/3$ along the direction of the beam. Relying on ones intuition, this implies that we may expect that the layer in a stellar or planetary atmosphere from which the radiation that we observe originates should be at an optical depth $\tau_\nu \simeq 2/3$, measured from outside in (so $\tau_\nu = 0$ is at the position of the observer). Later on in these lectures we will show that this is indeed the case.

Source function

The *source function* is defined as

$$S_\nu \equiv \frac{\eta_\nu}{\chi_\nu} \tag{4.29}$$

and has dimensions erg cm^{-2} sec^{-1} hz^{-1} sr^{-1}. It has the same units as the specific intensity, therefore S_ν and I_ν can be added and subtracted (see eq. 4.36). The extinction and emission coefficients are *local* quantities, implying the source function is independent of the adopted geometry.

The meaning of the source function can be understood as follows: The number of photons that is emitted per bandwidth and time interval from an elementary volume of cross-section dO and length ds into all directions is $N_{\rm em} = (4\pi/h\nu)\,\eta_\nu \, dO \, ds \, d\nu \, dt$, where we have assumed that the emission is isotropic. The factor 4π is the result from the integration over all solid

angles, and $h\nu$ converts energy to number of photons (see eq. 4.7). From the definition of optical depth, eq. (4.22), it follows that $\eta_\nu\, ds = (\eta_\nu/\chi_\nu)\,\chi_\nu\, ds = S_\nu\, d\tau_\nu$. So, we may also write for the number of emitted photons

$$N_{\rm em} = S_\nu\, d\tau_\nu\, \frac{4\pi}{h\nu}\, d\nu\, dt\, dO \tag{4.30}$$

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

Using an analogous reasoning we may derive the number of photons, coming from all directions, that is absorbed per unit optical depth. Using the definition of the extinction coefficient (eq. 4.3) we find that the number of photons in a ray traveling in direction $\mathbf{n}$ into solid angle $d\omega$ that suffer extinction is given by $dN_{\rm abs} = \chi_\nu(\mathbf{r})\, I_\nu(\mathbf{r},\mathbf{n})\, dO\, ds\, d\omega\, d\nu\, dt/h\nu$, where we have assumed that the extinction is isotropic. Integration over all solid angles, again using the definition of the optical depth, gives

$$N_{\rm abs} = J_\nu\, d\tau_\nu\, \frac{4\pi}{h\nu}\, d\nu\, dt\, dO \tag{4.31}$$

———

If several extinction/emission processes play a role the total source function is

$$S_\nu \equiv \frac{\eta_\nu}{\chi_\nu} = \frac{\sum_i \eta_{\nu,i}}{\sum_i \chi_{\nu,i}} \tag{4.32}$$

The subscript i labels all different processes contributing at frequency ν (see eq. 4.9).

Let us for a moment get ahead of things such that we can already give the proto-typical form of the source function. In a medium that features both absorption and scattering processes, Kirchhoff's law (eq.6.3) implies that for the thermal emission $\eta_\nu^{\rm th} = \kappa_\nu B_\nu$; if we assume isotropic and coherent (i.e. $\nu = \nu'$) continuum scattering, the emission term for scattering is given by $\eta_\nu^{\rm sc} = \sigma_\nu J_\nu$. For the proto-typical source function we get

$$S_\nu = \frac{\kappa_\nu B_\nu + \sigma_\nu J_\nu}{\kappa_\nu + \sigma_\nu} \tag{4.33}$$

In section 9.4 the proto-typical source function will be discussed in more detail.

The frequency integrated or *total source function* is

$$S \equiv \int_0^\infty S_\nu\, d\nu \tag{4.34}$$

4.4 Boundary conditions and formal solution of the planer transfer equation

We define the optical depth scale in a geometry of planar layers such that for z increasing in the direction to the observer, τ_ν is decreasing, i.e.

$$d\tau_\nu(z) = -\chi_\nu(z)dz \tag{4.35}$$

This implies that for an observer at $z = \infty$, the optical depth $\tau_\nu(\infty) = 0$. Adopting this optical depth scale, and using the source function, the planar transfer equation (eq. 4.18) may be written in its standard form

$$\mu\frac{\mathrm{d}I_\nu(\tau_\nu,\mu)}{\mathrm{d}\tau_\nu} = I_\nu(\tau_\nu,\mu) - S_\nu(\tau_\nu,\mu) \tag{4.36}$$

Note that we measure τ_ν along the z direction. The relation between the optical depth in the μ direction, $\tau_{\nu\mu}$, and that in the z direction (for which $\mu = 1$) is $\tau_{\nu\mu} = \tau_\nu/\mu$.

Boundary conditions in axial symmetry

Transfer equation (4.36) has an upper boundary condition for inward directed beams ($-1 \leq \mu \leq 0$) and a lower boundary condition for outward directed beams ($0 \leq \mu \leq 1$). How might these boundary conditions look like? If the outer boundary condition is dark space, the condition is

$$I_\nu(0,\mu) = 0 \qquad \text{for } -1 \leq \mu \leq 0. \tag{4.37}$$

It is quite customary to adopt this condition when modelling stars. For some stars though (e.g. Asymptotic Giant Branch stars), the mean interstellar radiation field $I_\nu^{\mathrm{ISRF}} = (c/4\pi)\,u_\nu$ is usually considered as upper boundary condition, where u_ν is the monochromatic interstellar energy density (see Eq.3.15). In that case

$$I_\nu(0,\mu) = J_\nu^{\mathrm{ISRF}}(0,\mu) \qquad \text{for } -1 \leq \mu \leq 0, \tag{4.38}$$

where the ISRF is assumed to be isotropic, i.e. $J_\nu^{\mathrm{ISRF}}(0,\mu) = J_\nu^{\mathrm{ISRF}}(0)$. In case one is modelling the surface of a planet illuminated by the host star the situation is more complex. Radiation is incident from above from (almost) a single direction, that of the star. Lets say the star is positioned at $(\mu_\circ, \phi_\circ)$. Simply look out of the window in daytime to see an illustration of this situation. Often (see Hansen 2008, ApJS 179, 484), this is ignored and the host star radiation is assumed to be incident uniformly from all incoming directions (so, Eq. 4.38 is adopted with the mean interstellar radiation field replaced by $J_\nu^\star = W(d)\,I_\nu^\star$, where $I_\nu^\star$ is the specific intensity of the host star). Alternatively, one assumes the condition

$$I_\nu(0,\mu) = I_\nu^\star(0,\mu_\circ) \qquad \text{for } -1 \leq \mu \leq 0, \tag{4.39}$$

for incoming starlight that – if axial symmetry is to be preserved – is essentially distributed over a ring in the sky (covering all azimuthal angles ϕ) at fixed polar angle $\mu_\circ$, and that is

emitting a specific intensity $I_\nu^\star$. This assumption is usually applied to describe the incident stellar flux assuming the incoming radiation to be plane parallel at the planet surface (see Eq. 12.11).

We will return to boundary conditions for exo-planets in Chapter 12.

The lower boundary condition specifies the specific intensity coming from the interior. In case of a semi-infinite atmosphere one usually applies the diffusion approximation (see Eq. 4.60)

$$I_\nu(\tau_{\max,\nu},\mu) = S_\nu(\tau_\nu) + \mu \left.\frac{\mathrm{d}S_\nu}{\mathrm{d}\tau_\nu}\right|_{\tau_{\max,\nu}} \qquad \text{for } 0 \le \mu \le 1, \tag{4.40}$$

for both stars and planets. Here, $\tau_{\max,\nu}$ is the optical depth at the lower boundary. For stars, the origin of the energy (specific intensity is photon energy after all) coming from deep layers is thermonuclear reactions in the core. For planets the energy has two possible origins. The first is from reprocessed absorbed host star radiation and the second is from interior energy arising from slow loss of residual gravitational energy and from radioactive decay.

Formal solution

For known source function S_ν, eq. (4.36) has a formal solution. To find this solution we bring I_ν to the left hand side, divide by μ and multiply both sides by the integrating factor $\exp(-\tau_\nu/\mu)$. This yields

$$\left[\frac{\mathrm{d}I_\nu}{\mathrm{d}\tau_\nu} - \frac{I_\nu}{\mu}\right] e^{-\tau_\nu/\mu} = \frac{\mathrm{d}}{\mathrm{d}\tau_\nu}(I_\nu e^{-\tau_\nu/\mu}) = -\frac{S_\nu}{\mu} e^{-\tau_\nu/\mu} \tag{4.41}$$

Integration from τ_1 to τ_2, and multiplying by $-\exp(\tau_1/\mu)$ results in

$$I_\nu(\tau_1,\mu) = I_\nu(\tau_2,\mu)\, e^{-(\tau_2-\tau_1)/\mu} + \int_{\tau_1}^{\tau_2} S_\nu(t_\nu)\, e^{-(t_\nu-\tau_1)/\mu} \frac{dt_\nu}{\mu} \tag{4.42}$$

This is the formal solution of the equation of transfer. We will discuss the meaning of the two terms in this solution in § 4.6 using some example problems. Here we only point out that the boundary conditions are represented by $I_\nu(\tau_2,\mu)$ for outward beams ($0 \le \mu \le 1$) and by $I_\nu(\tau_1,\mu)$ for inward beams ($-1 \le \mu \le 0$).

4.5 Moments of the transfer equation

We introduce the first two moments of the transfer equation. The angular moments (i.e. with respect to μ) of the transfer equation yield results of deep physical significance and great mathematical utility when constructing, for instance, models of stellar and planetary atmospheres.

Planar layers

The zero order moment of the equation of transfer in a planar medium can be derived from eq. (4.36)

$$\begin{aligned} \frac{1}{2}\int_{-1}^{+1} \mu \frac{\mathrm{d}I_\nu}{\mathrm{d}\tau_\nu}\, d\mu &= \frac{1}{2}\int_{-1}^{+1} I_\nu\, d\mu - \frac{1}{2}\int_{-1}^{+1} S_\nu\, d\mu \\ \frac{\mathrm{d}H_\nu}{\mathrm{d}\tau_\nu} &= J_\nu - S_\nu \end{aligned} \tag{4.43}$$

where we have assumed that the source function is isotropic. For the first order moment we find

$$\begin{aligned} \frac{1}{2}\int_{-1}^{+1} \mu^2 \frac{\mathrm{d}I_\nu}{\mathrm{d}\tau_\nu}\, d\mu &= \frac{1}{2}\int_{-1}^{+1} \mu I_\nu\, d\mu - \frac{1}{2}\int_{-1}^{+1} \mu S_\nu\, d\mu \\ \frac{\mathrm{d}K_\nu}{\mathrm{d}\tau_\nu} &= H_\nu \end{aligned} \tag{4.44}$$

as $\int_{-1}^{+1} \mu S_\nu\, d\mu = 0$ for an isotropic source function. Substituting eq. (4.44) in (4.43) then gives

$$\frac{\mathrm{d}^2 K_\nu}{\mathrm{d}\tau_\nu^2} = \frac{\mathrm{d}^2 (f_\nu J_\nu)}{\mathrm{d}\tau_\nu^2} = J_\nu - S_\nu \tag{4.45}$$

where we have used the Eddington factor defined in eq. (3.37). If f_ν and S_ν are known, the above equation reduces to a 2^{nd} order differential equation for the mean intensity. A method to solve this type of equation will be discussed in § 5.

Spherical shells

The zero order moment of the equation of transfer in a spherical symmetric medium can be derived from eq. (4.21)

$$\frac{1}{r^2}\frac{\partial (r^2 H_\nu)}{\partial r} = \eta_\nu - \chi_\nu J_\nu \tag{4.46}$$

where we have assumed that η_ν and χ_ν are isotropic. For the first order moment we find

$$\frac{\partial K_\nu}{\partial r} + \frac{1}{r}(3K_\nu - J_\nu) = \frac{\partial (f_\nu J_\nu)}{\partial r} + \frac{1}{r}(3f_\nu - 1)J_\nu = -\chi_\nu H_\nu \tag{4.47}$$

Similar to the case of a planar geometry, we may combine eq. (4.46) and (4.47) to arrive at a 2^{nd} order differential equation for the mean intensity.

4.6 Simple examples of transfer problems

We will now solve the equation of transfer for some simple cases.

Radiation from a homogeneous finite slab

Let us consider a homogeneous finite slab. Homogeneity implies that χ_ν and η_ν, therefore also S_ν, are constant. If the integrated optical depth in the normal direction of the slab is τ_ν the emerging intensity at $\tau_1 = 0$ is

$$I_\nu(0,\mu) = I_\nu(\tau_\nu,\mu)\, e^{-\tau_\nu/\mu} + S_\nu \left[1 - e^{-\tau_\nu/\mu}\right] \tag{4.48}$$

We concentrate on the normal direction $\mu = 1$ and for convenience drop reference to the angle dependence. This reduces eq. (4.48) to

$$I_\nu(0) = I_\nu(\tau_\nu)\, e^{-\tau_\nu} + S_\nu \left[1 - e^{-\tau_\nu}\right] \tag{4.49}$$

The first term on the right hand side describes the weakening of the radiation that is incident to the far side of the slab (viewed from the direction of the observer). The second term on the right hand side gives the contribution of radiation emitted by the slab itself. Let us analyse the two limiting cases of this solution.

In the optically thin limit ($\tau_\nu \ll 1$, such that $\exp(-\tau_\nu) \simeq 1 - \tau_\nu$) we find

$$I_\nu(0) \simeq I_\nu(\tau_\nu) + (S_\nu - I_\nu(\tau_\nu))\, \tau_\nu \tag{4.50}$$

If no radiation is incident at the far side ($I_\nu(\tau_\nu) = 0$) it follows that $I_\nu(0) \simeq S_\nu \tau_\nu$. This is to be expected as in the optically thin case we observe emission from almost all parts of the slab. The emerging radiation therefore must be $I_\nu(0) \simeq \eta_\nu D = S_\nu \chi_\nu D = S_\nu \tau_\nu$, where D is the geometrical thickness of the slab. If no radiation is emitted by the slab itself ($S_\nu = 0$) then $I_\nu(0) \approx I_\nu(\tau_\nu)$, which is obvious – one sees through the slab.

If the slab is optically thick ($\tau_\nu \gg 1$) then

$$I_\nu(0) \simeq S_\nu \tag{4.51}$$

The radiation $I_\nu(\tau_\nu)$ that is incident at the far side does not penetrate through the slab. One only observes the source function in the slab, irrespective of the nature of the extinction. The nature of the medium is only relevant for the source function S_ν.

We rewrite the solution of eq. (4.49) to

$$I_\nu(0) = I_\nu(\tau_\nu) + (S_\nu - I_\nu(\tau_\nu)) \left[1 - e^{-\tau_\nu}\right] \tag{4.52}$$

Note that the term $[1 - \exp(-\tau_\nu)]$ is always positive and that therefore $I_\nu(0) < I_\nu(\tau_\nu)$ if $S_\nu > I_\nu(\tau_\nu)$ (see figure 4.4). This is an important result. It shows the principle of the formation of absorption *casu quo* emission lines in a plane-parallel atmosphere. Let us identify the incident intensity $I_\nu(\tau_\nu)$ with the intensity that is emitted by the stellar or planetary continuum. The homogeneous layer corresponds to the stellar or planetary photosphere, i.e. the region in which the spectral lines are formed. Is the line source function $S_\nu < I_\nu(\tau_\nu)$ then an absorption line will form, i.e. $I_\nu(0) < I_\nu(\tau_\nu)$. Is the line source function $S_\nu > I_\nu(\tau_\nu)$ then an emission line will form, i.e. $I_\nu(0) > I_\nu(\tau_\nu)$. Is the line source function equal to the intensity emitted by the continuum, i.e. $S_\nu = I_\nu(\tau_\nu)$, then no line will form, i.e. $I_\nu(0) = I_\nu(\tau_\nu)$.

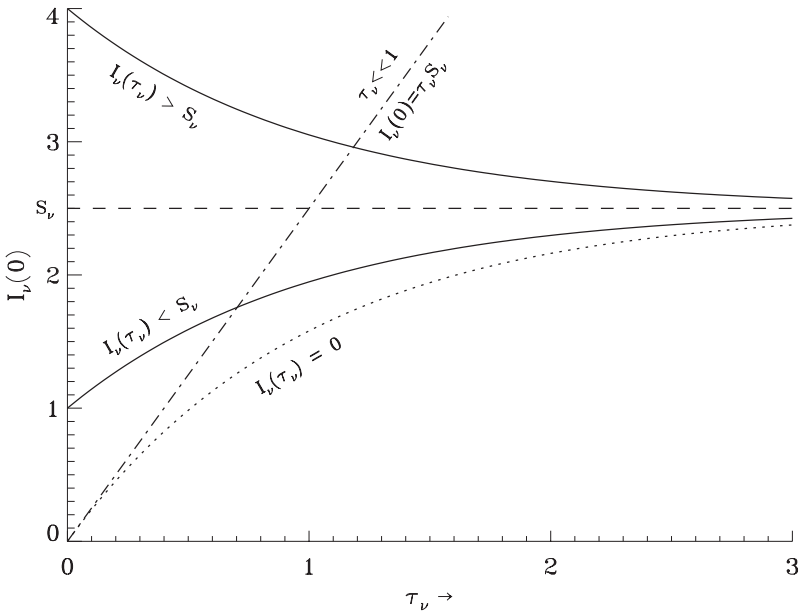


Figure 4.4: *The emerging specific intensity $I_\nu(0)$ for a homogeneous finite slab in which the optical depth in the normal direction is τ_ν. The emitted specific intensity approaches the source function S_ν from both sides, en reaches S_ν for a sufficiently large optical depth τ_ν of the slab.*

Radiation from a semi-infinite medium

The assumption of a homogeneous stellar atmosphere is not very realistic. To improve this model, we assume that the medium consists of multiple homogeneous planar layers, i.e. that variations occur in the z direction. As the density increases inward to very high values, the optical depth $\tau_2 \rightarrow \infty$. The emerging intensity ($\tau_1 = 0$) is then given by

$$I_\nu(0, \mu) = \int_0^\infty S_\nu(t_\nu)\, e^{-t_\nu/\mu}\, \frac{dt_\nu}{\mu} \tag{4.53}$$

Now let us assume that the source function is a linear function of optical depth, i.e.

$$S_\nu(\tau_\nu) = a_\nu + b_\nu \tau_\nu \tag{4.54}$$

Substitution in eq. (4.53) yields

$$I_\nu(0, \mu) = a_\nu + b_\nu\, \mu = S_\nu(\tau_\nu = \mu) \tag{4.55}$$

This is known as the *Eddington-Barbier approximation*. It shows that the emerging specific intensity in the outward normal direction ($\mu = 1$) is characterized by the source function at $\tau_\nu = 1$. This does *not* mean that all photons that are observed from this direction originate

from the layer where the optical depth $\tau_\nu = 1$. The effective contribution to the integrant $S_\nu \exp(-t_\nu)$ reaches over a broad range in optical depth, from the surface at $\tau_\nu = 0$ down to $\tau_\nu \approx 10$. From all of this part of the semi-infinite atmosphere photons will escape; *together* they are characterized by the value of the source function at $\tau_\nu = 1$.

For the emergent flux (see Eq. 3.21) one finds

$$\mathcal{F}_\nu^+ = 2\pi \int_0^1 I_\nu(0,\mu)\,\mu\,d\mu = \pi\, S_\nu(\tau_\nu = 2/3), \tag{4.56}$$

a result we will prove in one of the exercises at the end of this chapter. Notice that for a constant source function (i.e. $b_\nu = 0$, implying I_ν is isotropic) the emergent flux is given by $\mathcal{F}_\nu^+ = \pi I_\nu$, a result already obtained in Eq. (3.28).

The formation of spectral lines in a semi-infinite medium

As an example of the results that we have obtained above for the specific intensity in a semi-infinite medium we discuss the formation of spectral lines. A spectral line is always the result of a discrete (*bound-bound*) transition in an atom or ion, i.e. of a process that occurs in addition to continuous processes that take place at the frequency of interest.

The presence of a spectral lines has two consequences:

- The bound-bound process gives the opportunity for extinction, superposed on the continuous extinction. At the frequency of the spectral line the extinction coefficient of the medium will be larger than at frequencies next to the line.

- The source function associated with the bound-bound process may differ from that of continuous processes at the line frequency.

Together these two effects may cause that the spectral line is observable in the emerging intensity, either in emission or in absorption relative to the continuum background.

We examine the special case that the continuum and line source functions are equal, but may vary (together) with depth in the medium. The effect of the extra line extinction is that the layer of optical depth $\tau_\nu = 1$, which is representative of the emerging photons, is located further out for line frequencies relative to continuum frequencies in the spectrum next to the line. In other words: one can see less deep in the medium at line frequencies as at these frequencies additional bound-bound extinction processes take place. The four panels shown in figure 4.5 illustrate this.

Note that an optically thick medium with identical continuum and line source function will only show a spectral line if the source function changes with depth, such that a different sampling will cause a different emerging specific intensity. In the four panels the source

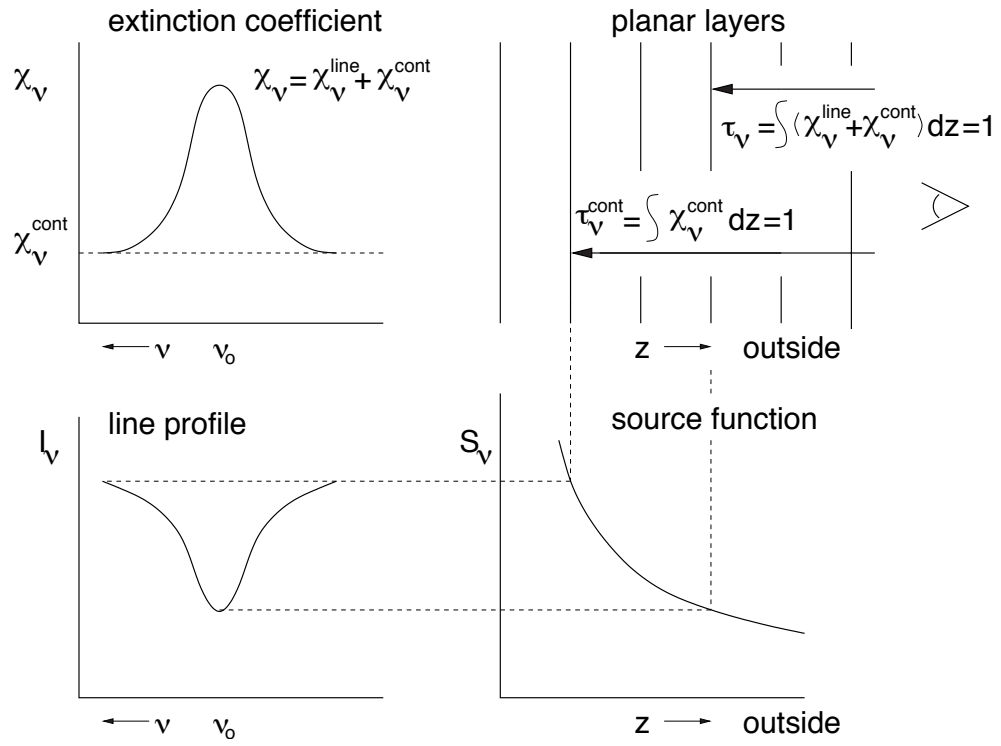


Figure 4.5: *Four panels illustrating line formation in a semi-infinite medium consisting of planar layers. The extinction coefficient χ_ν (top left) determines the location of $\tau_\nu = 1$ for each frequency ν (top right). As an example the location of line center is shown. This location defines the value of the source function (bottom right) to which the observed specific intensity will be equal to (bottom left). It is assumed that the continuum and line source function are identical (such that only one curve needs to be plotted in the bottom left panel). Because the source function decreases with increasing distance the spectral line is in absorption.*

function decreases outward such that spectral lines will *always* show a lower specific intensity compared to the continuum: they will be absorption lines.

The effect of line extinction on the characteristic depth of formation $\tau_\nu = 1$ can be very large as the bound-bound extinction coefficient at line centre is often orders of magnitudes larger than the continuum extinction coefficient. A "strong" spectral line may thus sample the emerging intensity in a geometrically thick layer. This implies that the shape of an observed strong spectral line does not so much reflect the behaviour of the extinction coefficient as a function of frequency, rather it reflects the run of the source function with depth in the atmosphere, sampled by the extinction coefficient. Throughout the line profile it is the depth one views into the atmosphere that is changing, and by that the value of the source function that we see.

The above discussion is not limited to rapid changes of the extinction coefficient in small

frequency intervals, as is the case for bound-bound processes. It holds equally well for widely separated frequencies that measure the slow change of the continuum extinction coefficient. Although in this case the assumption that the two source functions are identical is usually not valid, and the extinction coefficients thus measure their own source function.

Radiation deep inside a semi-infinite medium

The above discussed Eddington-Barbier approximation describes the radiation field close to the surface. Now we will direct our attention to the behaviour of the radiation quantities at great optical depth, $\tau_\nu \gg 1$, i.e. deep inside the medium, far below the surface. Here all characteristic length scales will be larger than the mean free path of the photons, therefore the particles will "see" an almost isotropic radiation field. As large optical depths are reached in regions of high density, collisional processes, i.e. real absorptions and emissions, will dominate over photon scattering processes. Consequently, the properties of the medium will be close to that of thermodynamic equilibrium (see chapter 6) such that LTE is a good approximation. We formulate the source function as a Taylor-McLaurin expansion

$$S_\nu(t_\nu) = \sum_{n=0}^{\infty} \frac{(t_\nu - \tau_\nu)^n}{n!} \left[\frac{\mathrm{d}^n S_\nu(t_\nu)}{\mathrm{d}t_\nu^n}\right]_{\tau_\nu} \tag{4.57}$$

Substitution in eq. (4.53) (you may find it more insightful to use the rightmost term in eq. 4.42) gives for an outward directed beam ($0 \leq \mu \leq 1$)

$$\begin{aligned} I_\nu(t_\nu,\mu) &= \sum_{n=0}^{\infty} \frac{1}{n!} \left[\frac{\mathrm{d}^n S_\nu(t_\nu)}{\mathrm{d}t_\nu^n}\right]_{\tau_\nu} \int_{\tau_\nu}^{\infty} (t_\nu - \tau_\nu)^n e^{-(t_\nu-\tau_\nu)/\mu}\, dt_\nu/\mu \\ &= \sum_{n=0}^{\infty} \left[\frac{\mathrm{d}^n S_\nu(t_\nu)}{\mathrm{d}t_\nu^n}\right]_{\tau_\nu} \frac{1}{n!} \int_0^{\infty} x^n e^{-x/\mu}\, dx/\mu \\ &= \sum_{n=0}^{\infty} \mu^n \left[\frac{\mathrm{d}^n S_\nu(t_\nu)}{\mathrm{d}t_\nu^n}\right]_{\tau_\nu} = S_\nu(\tau_\nu) + \mu \left.\frac{\mathrm{d}S_\nu}{\mathrm{d}t_\nu}\right|_{\tau_\nu} + \mu^2 \left.\frac{\mathrm{d}^2 S_\nu}{\mathrm{d}t_\nu^2}\right|_{\tau_\nu} + \cdots \end{aligned} \tag{4.58}$$

where we have made use of eq. (4.70). For an inward directed beam ($-1 \leq \mu \leq 0$) one recovers an identical expression save for correction terms of order $e^{-\tau_\nu/\mu}$. As we consider large optical depths these corrections are unimportant. We may therefore apply the above expression for radiation in all directions; moreover it applies for small optical depths as well, provided that $\mu > 0$.

Using eq. (3.20) we find for the net flux

$$\mathcal{F}_\nu = \sum_{n=0}^{\infty} \frac{4\pi}{2n+3} \left[\frac{\mathrm{d}^{2n+1} S_\nu(t_\nu)}{\mathrm{d}t_\nu^{2n+1}}\right]_{\tau_\nu} = \frac{4\pi}{3} \left.\frac{\mathrm{d}S_\nu}{\mathrm{d}t_\nu}\right|_{\tau_\nu} + \frac{4\pi}{5} \left.\frac{\mathrm{d}^3 S_\nu}{\mathrm{d}t_\nu^3}\right|_{\tau_\nu} + \cdots \tag{4.59}$$

To get a rough idea of the rate of convergence of these expansions we make the simple order of magnitude estimate $|\mathrm{d}^n S_\nu/\mathrm{d}\tau_\nu^n| \sim S_\nu/\tau_\nu^n$, such that the ratio between the differentials

$\left|\mathrm{d}^{n+2}S_\nu/\mathrm{d}\tau_\nu^{n+2}\right| / \left|\mathrm{d}^n S_\nu/\mathrm{d}\tau_\nu^n\right| \sim 1/\tau_\nu^2$. This guarantees fast convergence and yields for $\tau_\nu \gg 1$ the following approximations

$$I_\nu(\tau_\nu,\mu) \simeq S_\nu(\tau_\nu) + \mu \left.\frac{\mathrm{d}S_\nu}{\mathrm{d}t_\nu}\right|_{\tau_\nu} \tag{4.60}$$

$$J_\nu(\tau_\nu) \simeq S_\nu(\tau_\nu) \tag{4.61}$$

$$\mathcal{F}_\nu(\tau_\nu) \simeq \frac{4\pi}{3} \left.\frac{\mathrm{d}S_\nu}{\mathrm{d}t_\nu}\right|_{\tau_\nu} \tag{4.62}$$

and

$$K_\nu(\tau_\nu) \simeq \frac{1}{3} S_\nu(\tau_\nu) \tag{4.63}$$

where we have also given the end results for the mean intensity J_ν and the second order moment K_ν for completeness. In the expansion for the specific intensity we also take the first order term into account because we want to have a non-zero value for the flux. The isotropic component of the mean intensity is determined by the value of the source function. The net flux is determined by the anisotropic component of the source function. This is logical as $\mathcal{F}_\nu$ measures the difference between $\mathcal{F}_\nu^+$ and $\mathcal{F}_\nu^-$. For the radial component only $I_\nu^+ \simeq S_\nu + \mathrm{d}S_\nu/\mathrm{d}\tau_\nu$ and $I_\nu^- \simeq S_\nu - \mathrm{d}S_\nu/\mathrm{d}\tau_\nu$. The difference (relevant for the net flux) is proportional to $\mathrm{d}S_\nu/\mathrm{d}\tau_\nu$. This implies that in order to transport radiation outward the source function has to increase inward. In LTE this corresponds to a temperature that increases in the inward direction (see § 6.7).

Note that

$$\lim_{\tau_\nu\to\infty} \frac{K_\nu(\tau_\nu)}{J_\nu(\tau_\nu)} = \lim_{\tau_\nu\to\infty} f_\nu(\tau_\nu) = \frac{1}{3} \tag{4.64}$$

a result that we already obtained in § 3.6 for a strictly isotropic radiation field. Equation (4.64) therefore expresses that for ever increasing optical depth the radiation field becomes more and more isotropic. This can also be learned from a simple order of magnitude estimate of the anisotropy of the specific intensity $\left|\mathrm{d}S_\nu/\mathrm{d}\tau_\nu\right| / S_\nu \sim 1/\tau_\nu$, i.e. if τ_ν increases the anisotropy decreases.

Diffusion approximation

At large optical depth the properties of the medium approach (local) thermodynamic equilibrium. We discuss the specifics of these equilibria in more detail in Chapter 6. For now we mention that for thermodynamic equilibrium $S_\nu = B_\nu$ (see § 6.2). For the flux this implies (see eq. 4.62)

$$\mathcal{F}_\nu \simeq \frac{4\pi}{3}\frac{\partial B_\nu}{\partial \tau_\nu} = -\frac{4\pi}{3}\frac{1}{\chi_\nu}\frac{\partial B_\nu}{\partial z} = -\frac{4\pi}{3}\frac{1}{\chi_\nu}\frac{\partial B_\nu}{\partial T}\frac{\mathrm{d}T}{\mathrm{d}z} \tag{4.65}$$

This equation is known as the *diffusion approximation* because of the similarity of the equations to that of (other) diffusion processes that have a form such that the transported flux of a

quantity is the product of a diffusion coefficient times the gradient of the quantity itself. The coefficient $(4\pi/3\chi_\nu)\mathrm{d}B_\nu/\mathrm{d}T$ is also referred to as the *radiative thermal conductivity*.

In those parts of the atmosphere that can not be observed – or more general in the stellar interior – the total flux $\mathcal{F}$ is a much more interesting quantity than is the monochromatic flux $\mathcal{F}_\nu$. If we integrate eq. (4.65) over frequency and introduce the *Rosseland mean extinction coefficient* χ_R (dimension cm^{-1}), defined as

$$\frac{1}{\chi_\mathrm{R}} \equiv \frac{\int_0^\infty (1/\chi_\nu)(\partial B_\nu/\partial T)\, d\nu}{\int_0^\infty (\partial B_\nu/\partial T)\, d\nu} \tag{4.66}$$

we find using (see eq. 6.11)

$$\int_0^\infty \frac{\partial B_\nu}{\partial T}\, d\nu = \frac{\partial}{\partial T}\int_0^\infty B_\nu\, d\nu = \frac{\partial B}{\partial T} = \frac{\partial}{\partial T}\left(\frac{\sigma}{\pi}T^4\right) = \frac{4\sigma}{\pi}T^3 \tag{4.67}$$

for the total flux

$$\mathcal{F} = -\frac{16}{3}\frac{\sigma T^3}{\chi_\mathrm{R}}\frac{\mathrm{d}T}{\mathrm{d}z} \tag{4.68}$$

where $16\sigma T^3/3\chi_\mathrm{R}$ is the *effective radiative thermal conductivity*. This equation shows that in LTE a net outward radiative flux is accompanied by an inward increase of the temperature.

The general form of the above diffusion equation is given by

$$\boldsymbol{\mathcal{F}}(\mathbf{r},t) = \frac{16}{3}\frac{\sigma T^3}{\chi_\mathrm{R}}\nabla T = \frac{16}{3}\frac{\sigma T^3}{\chi_\mathrm{R}}\left(\frac{\partial T}{\partial x}, \frac{\partial T}{\partial y}, \frac{\partial T}{\partial z}\right) \tag{4.69}$$

A formal derivation of this equation is given in Wehrse, Baschek, & von Waldenfels (2000)

We will return to the diffusion approximation in chapter 10.

Exercise 4.1

This could be a nice exam question. Suppose that in Sherwood Forest, the average radius of a tree is $R = 1$ m and that the average number of trees per unit area is $\Sigma = 0.005\,\mathrm{m}^{-2}$. When Robin Hood or Lady Marian shoots an arrow, it flies horizontally until it strikes a tree.

a) What is the mean cross section of a tree in m?

b) If Robin Hood shoots an arrow in a random direction, how far, on average, will it travel before it strikes a tree?

c) If Lady Marian shoots a total of 1000 arrows in random directions, how many, on average, will travel at least 500 m?

Exercise 4.2

Show that for a medium in which no photons are created, destructed or converted, but in which only pure scattering processes occur, the source function is equal to the mean intensity: $S_\nu = J_\nu$. The importance of this result is that in a scattering-dominated medium the source function at some position $\mathbf{r}$ may be the result of a radiation field that is generated far, far away, at some $\mathbf{r}\prime$ where conditions (temperature, density) are completely different from those at $\mathbf{r}$.

Exercise 4.3

Derive the moment equations (4.46) and (4.47) in case of a spherically symmetric medium. These equations play an important role in the modelling of, for instance, stellar winds, supernova remnants, and Bok globules.

Exercise 4.4

Stellar light passes through an interstellar cloud. The continuum specific intensity of the starlight is $I_\circ = 1$, independent of frequency. Spectral line absorption occurs in the cloud. The optical depth in the spectral line is given by $\tau(x) = \tau_\circ \phi(x)$, where $\phi(x)$ is a 'bell shaped' function given by Eq. 13.20. The parameter x is a dimensionless measure of frequency. Use a spreadsheet program to create a graph of the spectrum of the emerging specific intensity I_x in the range $x = [-4.0, +4, 0]$, for $\tau_\circ$ = 0.1, 1.0, and 10.

Exercise 4.5

a) Give the emerging specific intensity, as a function of beam direction μ, for a homogeneous semi-infinite slab. The result tells you that for these assumptions, no matter the angle at which you view the surface of the slab, you always see the same surface brightness.

b) Give the specific intensity in a direction $\mu > 0$ inside of a homogeneous, semi-infinite medium.

Exercise 4.6

A radio astronomer states that the radio intensity that she observes from the center of a spherical looking interstellar cloud of diameter D is given by $I_\nu = \chi_\nu S_\nu D$. Give the five assumptions that she has made.

Exercise 4.7

a) Give the derivation of the Eddington-Barbier approximation (Eq. 4.55). This result is fundamental in understanding spectral line formation.

b) Give the emerging specific intensity in case $S_\nu(\tau_\nu) = \sum_{n=0}^{\infty} a_{n,\nu}\tau_\nu^n$. Remember that

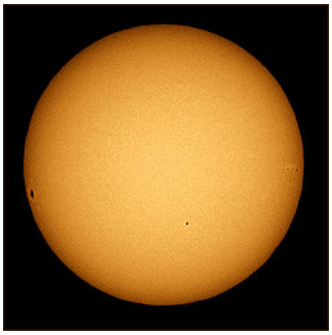
$$\int_0^\infty x^n e^{-x}\, dx = n! \tag{4.70}$$

Show that one recovers the Eddington-Barbier approximation if $a_{n,\nu} = 0 \; \forall \; n \geq 2$.

Exercise 4.8

a) The specific intensity of the solar disk in visual light declines from the center of the disk towards the edge. What does this tell about the variation of the source function as a function of height in the solar atmosphere?

b) This image of the solar disk is taken during the transit of Mercury on November 8, 2006. Locate Mercury.



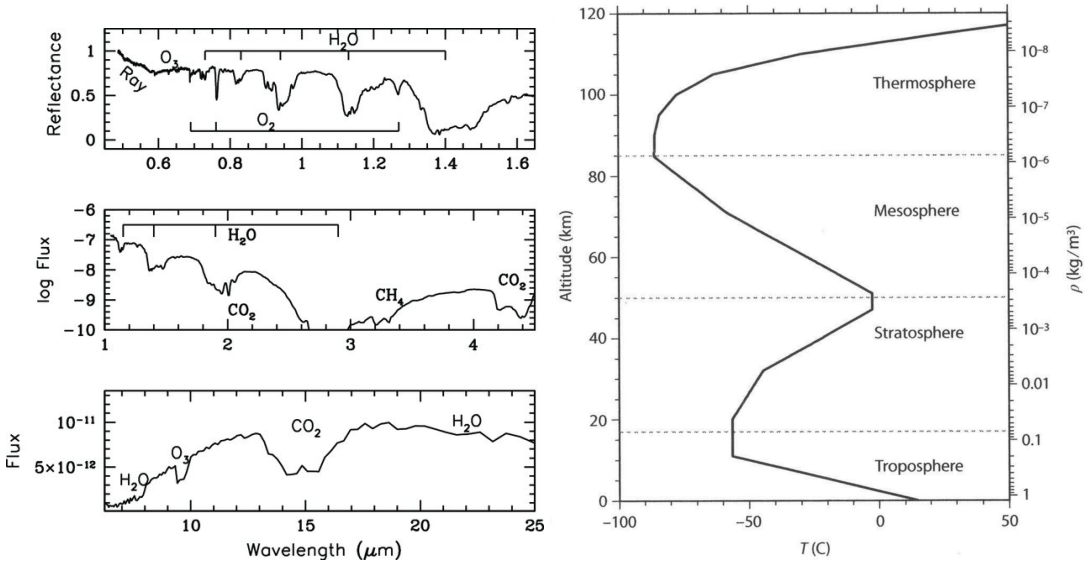


Figure 4.6: *Left panel: Earth's hemispherically averaged spectrum, i.e. Earth as an exoplanet. Top: Earth's visible spectrum from Earthshine measurements plotted as normalised reflectance (Turnbull et al. 2006). Middle: near-IR spectrum from NASA's EPOXI mission with flux in units of* $W\,m^{-2}\,lm^{-1}$ *(Robinson et al. 2011). Bottom: Earth's mid-IR spectrum as observed by Mars Global Surveyor en route to Mars with flux in units of* $W\,m^{-2}\,Hz^{-1}$ *(Christensen & Pearl, 1997). Major molecular absorption features are noted including Rayleigh scattering (noted as 'Ray'). Only strongly absorbing, globally mixed molecules are detectable. Right panel: the vertical structure of Earth's atmosphere typical of midlatitudes. From the 1976 U.S. Standard Atmosphere.*

Exercise 4.9

Figure 4.6 shows Earth's mid-infrared spectrum as seen from space and Earth's vertical temperature structure typical of mid-latitudes. From the ground up to about 60 km our atmosphere is in LTE (we will get to this in Sect. 6.7) and we adopt $S_\nu(z) = B_\nu(T(z))$ for the source function.

a) In which layer of Earth's atmosphere is most of the gas mass concentrated? It is fair to assume that the bulk of the spectral features seen in the figure are formed in this layer.

b) We focus on the CO_2 feature at $13-17\,\mu$m. The overal shape of the CO_2 profile is that of an absorption line. Explain why this is to be expected.

c) Notice a narrow emission feature (at about $15\,\mu$m) in the center of the CO_2 absorption. Explain the origin of this feature.

Exercise 4.10

Prove that in the Eddington Barbier approximation the emergent flux is given by Eq. 4.56.

Exercise 4.11

We assume a semi-infinite stellar atmosphere. The continuum extinction $\chi^{\rm cont}$ in the part of the spectrum that we are probing is independent of frequency and height. The source function is a linear function of optical depth. Using the continuum optical depth as our measure of vertical height, the source function is $S(\tau^{\rm cont}) = a + b\,\tau^{\rm cont}$. The spectrum contains a spectral line that has a source function that is the same as that of the continuum at every position in the atmosphere. The line extinction coefficient is given by $\chi_x^{\rm line} = \chi_\circ^{\rm line}\,\phi(x)$, where $\phi(x)$ is a 'bell shaped' function given by Eq. 13.20 and $\chi_\circ^{\rm line}$ is independent of height. The parameter x is a measure of frequency. The total optical depth $\tau_x = \tau^{\rm cont} + \tau_x^{\rm line}$. As the Eddington-Barbier approximation is valid, the emerging specific intensity in the normal direction is given by $I_x(0) = S(\tau_x = 1)$.

a) Show that $\tau_x = 1$ is reached for $\tau^{\rm cont} = 1/(1+q)$, where $q = \chi_x^{\rm line}/\chi^{\rm cont}$ is the ratio between the line extinction coefficient and continuum extinction coefficient.

b) Adopt $\chi_\circ^{\rm line} = 10$ and $\chi_{\rm circ}^{\rm cont} = 1$. Use a spreadsheet program to create a graph of the spectrum of the emerging specific intensity I_x in the range $x = [-4.0, +4.0]$, for (a, b) = (1.0, 0.0), (0.6, 0.4), and (0.2, 0.8). Note that a steeper dependence of the source function with optical depth produces a deeper absorption line.

Exercise 4.12

Make a schematic drawing similar to Fig. 4.5 for the case that the source function increases in the outward direction. What type of spectral line does this give?

Exercise 4.13

Show that eq. (4.58) is also valid for inward directed beams, i.e. for $-1 \leq \mu \leq 0$. Your professor agrees that this is quite heavy stuff. It is just to let you know that if you ever find yourself deep down in a star, move in the direction from where the surface brightness that you see is lowest to find the exit.

Exercise 4.14

a) Show that for the mean intensity a Taylor-McLaurin expansion of the source function (eq.4.57) yields

$$J_\nu(\tau_\nu) = \sum_{n=0}^{\infty} \frac{1}{2n+1} \left[\frac{\mathrm{d}^{2n} S_\nu(t_\nu)}{\mathrm{d}t_\nu^{2n}} \right]_{\tau_\nu} = S_\nu(\tau_\nu) + \frac{1}{3} \left. \frac{\mathrm{d}^2 S_\nu}{\mathrm{d}t_\nu^2} \right|_{\tau_\nu} + \cdots \tag{4.71}$$

b) Show that the second order moment is given by

$$K_\nu(\tau_\nu) = \sum_{n=0}^{\infty} \frac{1}{2n+3} \left[\frac{\mathrm{d}^{2n} S_\nu(t_\nu)}{\mathrm{d}t_\nu^{2n}} \right]_{\tau_\nu} = \frac{1}{3} S_\nu(\tau_\nu) + \frac{1}{5} \left. \frac{\mathrm{d}^2 S_\nu}{\mathrm{d}t_\nu^2} \right|_{\tau_\nu} + \cdots \tag{4.72}$$

c) What is the Eddington factor f_ν (Eq. 3.37) deep down in a star, where the (value of the) gradient in the source function is very small compared to the (value of the) source function itself.

Exercise 4.15

This could be a nice exam question. Consider a spherical homogeneous cloud of small dust particles. The cloud has a mass m and a radius $R = 0.1$ AU (1 AU = 1.5×10^{13} cm). The dust particles have a temperature T and radiate according to the Planck-function. The optical depth τ, as measured from the edge of the cloud to the center, is independent of frequency ('gray').

a) Give the specific intensity, mean intensity, and flux at the center of the cloud.

The dust particles have a frequency independent ('gray') extinction coefficient $\chi' = 100\,\mathrm{cm}^2/\mathrm{gr}$.

b) Derive an expression that gives the gray radial optical depth as a function of the basic parameters of the cloud.

c) What should be the minimum mass of the cloud to assure that the cloud is optically thick (i.e. $\tau \geq 1$)?

d) If we increase the radius of the cloud by a factor of two, but keep its mass the same, does this lead to a change in the specific intensity at the center (yes/no). Does the mean intensity at the center change (yes/no)? Does the flux at the center change (yes/no)?

The cloud represents the halo of a comet. This halo develops when the comet gets so close to the sun that because of heating small dust particles come off its surface layers. A comet has a typical initial diameter of 10 km and consists of material that has a typical density of $1\,\mathrm{gr\,cm}^{-3}$.

e) Will the halo of this comet be optically thick or optically thin?

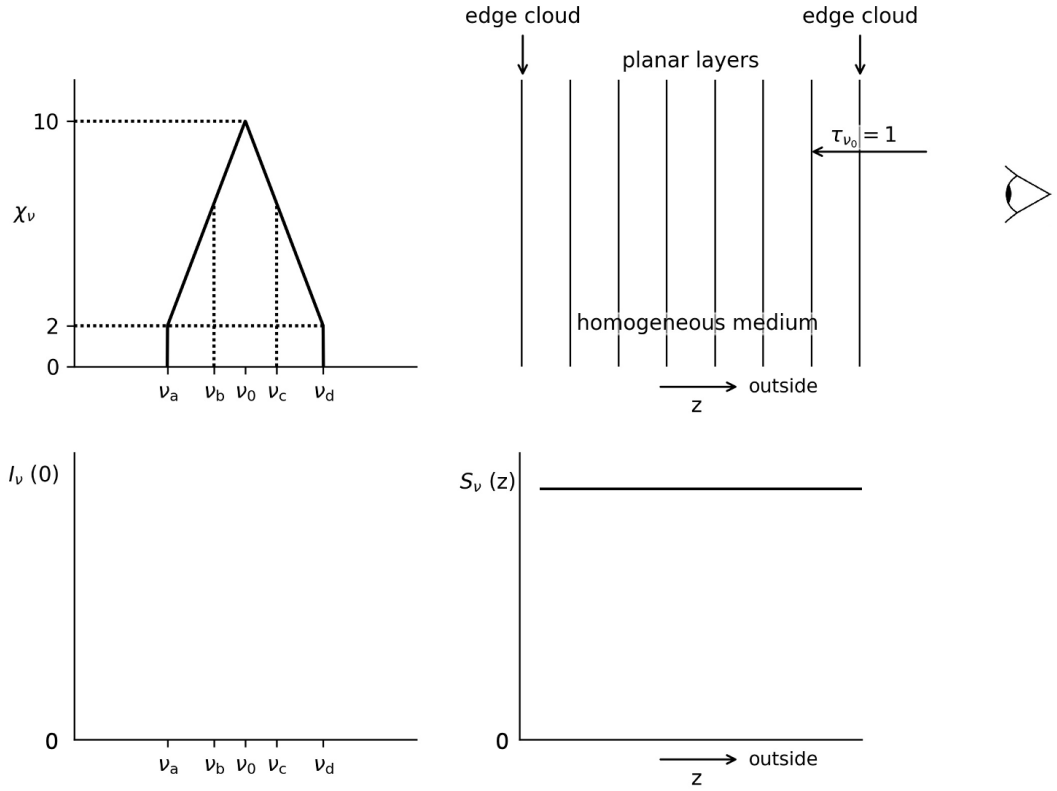


Figure 4.7: *Four-panel diagram for a spectral line formed in a homogeneous interstellar cloud.*

Exercise 4.16

This could be a nice exam question. We use the four-panel diagram (introduced in Fig. 4.5) to draw the shape of a spectral line that is formed in an interstellar cloud. For this purpose an empty diagram is given in Fig. 4.7). The cloud is spatially resolved and can be modelled by a homogeneous slab, i.e., by a geometry of planar layers. The background behind the cloud is dark, i.e., there is no radiation illuminating the cloud at the far side. As the slab is homogeneous, the density and temperature are constant as well as the source function of the spectral line we're considering. The spectral line is very optically thick at all line frequencies $[\nu_a, \nu_b]$. In the top left panel the extinction coefficient is given in arbitrary units. In all panels the axes are on a linear scale.

a) Which condition, in general, determines whether a line is in emission or in absorption?

b) Use the panel to reconstruct the line profile (lower left panel) that the cloud is emitting. In doing so, draw – in any case – the situation for frequencies $\nu_\mathrm{a}, \nu_\mathrm{o}$ and ν_d in the panels. Do so as quantitative (notably in the top right panel) and accurate as possible. Draw dashed lines from panel to panel to clarify your reasoning.

We now assume that the line is very optically thick only for frequencies in between $[\nu_{\rm b}, \nu_{\rm c}]$. For frequencies outside of this regime the line becomes optically thin. At $\nu_{\rm a}$ and $\nu_{\rm d}$ the optical depth $\tau_{\nu_{\rm a}} = \tau_{\nu_{\rm d}} = 0.7$.

c) Draw/sketch the line profile for this situation (add it to the lower left panel); do so quantitatively for the frequency points for which you know what the emerging intensity $I_\nu(0)$ should be and sketch the situation in between these points.

Exercise 4.17

In section 9.4 we will discuss in more detail that in a planar atmosphere the total flux $\mathcal{F}$ (see 4.68) is constant. The value for the constant is $\sigma\, T_{\rm eff}^4$, see eq. (3.23). Introduce, using eq. (4.35), the Rosseland optical depth $d\tau_{\rm Ross} = -\chi_{\rm R}\, dz$ and show that in the diffusion limit the temperature structure is given by

$$T^4(\tau_{\rm Ross}) \;=\; \frac{3}{4}\, T_{\rm eff}^4 \left(\tau_{\rm Ross} + \frac{2}{3}\right)$$

if one assumes that the diffusion approximation is valid down to $\tau_{\rm Ross} = 2/3$ and that the temperature T at $\tau_{\rm Ross} = 2/3$ is equal to $T_{\rm eff}$.

5

Numerical methods for solving the equation of transfer

In many cases the equation of transfer can not be solved analytically. In this chapter we discuss a number of numerical techniques that allow to compute the formal solution eq. (4.42). "Formal" implies that both S_ν and τ_ν are known functions. In principle, we may simply replace the integral over optical depth by a quadrature sum, such as for instance the *trapezium rule*

$$\int_{x_1}^{x_2} f(x)\, dx = \Delta x \left[\frac{1}{2} f_1 + \frac{1}{2} f_2\right] \tag{5.1}$$

where $f(x) = S(x)e^{-x}$. Solving this problem is trivial. In practice, however, this is not the way in which it is done. The reason for it is that this simple numerical quadrature is *a)* inaccurate because of the stongly non-linear (read: exponential) behaviour of the integrant $f(x)$, and *b)* very inefficient in terms of computation speed. The latter is so because the integrant contains an exponential function, which is very "expensive" to compute. In many complex models, especially those aimed at quantitative spectroscopy, the overal computation time is dominated by the time it costs to perform formal solutions (for all frequencies and angles at all grid points). Highly efficient numerical schemes are therefore indispensable.

One may distinguish between two classes of methods, those based on the *i)* integral solution (eq. 4.42), or *ii)* differential solution (eq. 4.36) of the equation of transfer. The latter class can again be divided in a *a)* first-order form, or *b)* a second-order form. The *formal solution using short characteristics* uses the integral solution. We will first discuss this method. Then we will focus on a second-order differential method that is refered to as *Feautrier method*, named after the person that devised this scheme. Until recently the accurate and efficient Feautrier method (which does not feature exponential functions) has been the prefered numerical tool for solving the transfer equation. Presently, there is a renewed interest in the method of short characteristics. This is so because of the relative ease with which this formalism can be implemented in 2D and 3D transfer problems. For both the 1D and multiD-problems there is a recent increasing interest in an ingeneous adaptation of the *discontinous finite element* method, which is based on the first-order differential form. We will not discuss this last method.

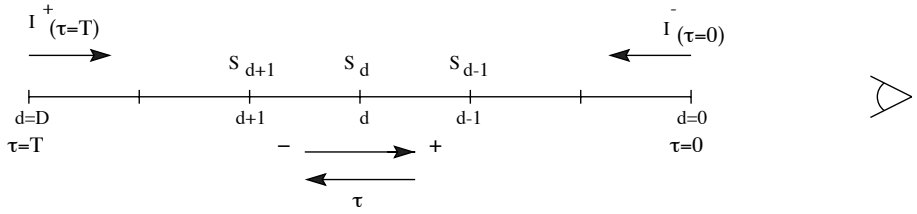


Figure 5.1: *Specification of the grid points in our description of the method of short characteristics.*

5.1 Short characteristics

The *method of short characteristics* uses the integral solution eq. (4.42) of the equation of transfer. Let us consider a finite slab in planar geometry, through which the total optical depth of the medium in the normal direction is T_ν. Along lines of sight in the outward direction, i.e. for $0 \leq \mu \leq 1$, or, phrased differently, in the direction in which the optical depth decreases, we have

$$I_\nu^+(\tau_\nu, \mu) = I_\nu^+(T_\nu, \mu)e^{-(T_\nu - \tau_\nu)/\mu} + \int_{T_\nu}^{\tau_\nu} S_\nu(t_\nu)e^{-(t_\nu - \tau_\nu)/\mu}\frac{dt_\nu}{\mu} \tag{5.2}$$

where $I_\nu^+(T_\nu, \mu)$ is the incident intensity (in the + direction) at $\tau_\nu = T_\nu$. Similarly, for an inward directed beam, i.e. $-1 \leq \mu \leq 0$, we find using the boundary condition $I_\nu^-(0, \mu)$ that

$$I_\nu^-(\tau_\nu, \mu) = I_\nu^-(0, \mu)e^{\tau_\nu/\mu} + \int_0^{\tau_\nu} S_\nu(t_\nu)e^{-(\tau_\nu - t_\nu)/\mu}\frac{dt_\nu}{(-\mu)} \tag{5.3}$$

Perhaps, intuitively, this last solution may seem somewhat odd. However, realize that μ is negative.

We choose a set of optical depth points $\{\tau_d\}, d = 0, ..., D$ where $\tau_0 = 0$ and $\tau_D = T$. For convenience we drop the frequency subscript. The boundary conditions are given by $I_D^+(\mu, \nu) = I^+(T, \mu, \nu)$ and $I_0^-(\mu, \nu) = I^-(0, \mu, \nu)$. Figure 5.1 illustrates this grid specification.

The above equations provide the solution along beams that pass through the slab at an angle θ, where $\mu = \cos\theta$. We now write the solution in a local form, such that it appears that the specific intensity at point d is only a function of the source function and the boundary conditions $I^+(\tau_{d+1}, \mu, \nu)$ and $I^-(\tau_{d-1}, \mu, \nu)$. This yields

$$I^+(\tau_d, \mu, \nu) = I^+(\tau_{d+1}, \mu, \nu)\, e^{-\Delta\tau_{d+\frac{1}{2}}} + \Delta I_d^+(S, \mu, \nu) \tag{5.4}$$

and

$$I^-(\tau_d, \mu, \nu) = I^-(\tau_{d-1}, \mu, \nu)\, e^{-\Delta\tau_{d-\frac{1}{2}}} + \Delta I_d^-(S, \mu, \nu) \tag{5.5}$$

where

$$\Delta\tau_{d+\frac{1}{2}} = (\tau_{d+1} - \tau_d)/\left|\mu\right| \tag{5.6}$$

The increase in intensity, $\Delta I_d^{\pm}$, is the contribution of the source function integrated over the grid interval. We now assume that locally (over the range $\{\tau_{d-1}, \tau_{d+1}\}$) the source function may be approximated, for instance by first- or second-order polynomials (Kunasz & Auer 1988; van Noort et al. 2002), Bézier curves (Hayek et al. 2010; Holzreuter & Solanki 2012; Auer 2003; Hennicker et al. 2020) or cubic Hermite splines (Ibgui et al. 2013). Let us here adopt a parabolic function, i.e. second-order polynomial

$$S(\tau) = a + b\tau + c\tau^2 \tag{5.7}$$

Note that if $c = 0$ this expression reduces to the Eddington-Barbier approximation (see § 4.6), which assumes that the source function is linear in each grid interval $\{\tau_d, \tau_{d+1}\}$. Using the adopted behaviour (5.7) we may analytically solve the transfer equation over the relevant grid interval, and we may express this solution as

$$\Delta I_d^{\pm} = \alpha_d^{\pm} S_{d-1} + \beta_d^{\pm} S_d + \gamma_d^{\pm} S_{d+1} \tag{5.8}$$

After a heroic amount of algebra and integration we find for the parabolic interpolation coefficients

$$\begin{aligned}
\alpha_d^- &= e_{0d} + \left[e_{2d} - \left(\Delta\tau_{d+\frac{1}{2}} + 2\Delta\tau_{d-\frac{1}{2}}\right) e_{1d}\right] / \left[\Delta\tau_{d-\frac{1}{2}} \left(\Delta\tau_{d+\frac{1}{2}} + \Delta\tau_{d-\frac{1}{2}}\right)\right] \\
\beta_d^- &= \left[\left(\Delta\tau_{d+\frac{1}{2}} + \Delta\tau_{d-\frac{1}{2}}\right) e_{1d} - e_{2d}\right] / \left[\Delta\tau_{d-\frac{1}{2}}\Delta\tau_{d+\frac{1}{2}}\right] \\
\gamma_d^- &= \left[e_{2d} - \Delta\tau_{d-\frac{1}{2}} e_{1d}\right] / \left[\Delta\tau_{d+\frac{1}{2}} \left(\Delta\tau_{d+\frac{1}{2}} + \Delta\tau_{d-\frac{1}{2}}\right)\right] \\
\alpha_d^+ &= \left[e_{2d+1} - \Delta\tau_{d+\frac{1}{2}} e_{1d+1}\right] / \left[\Delta\tau_{d-\frac{1}{2}} \left(\Delta\tau_{d+\frac{1}{2}} + \Delta\tau_{d-\frac{1}{2}}\right)\right] \\
\beta_d^+ &= \left[\left(\Delta\tau_{d+\frac{1}{2}} + \Delta\tau_{d-\frac{1}{2}}\right) e_{1d+1} - e_{2d+1}\right] / \left[\Delta\tau_{d-\frac{1}{2}}\Delta\tau_{d+\frac{1}{2}}\right] \\
\gamma_d^+ &= e_{0d+1} + \\
&\quad \left[e_{2d+1} - \left(\Delta\tau_{d-\frac{1}{2}} + 2\Delta\tau_{d+\frac{1}{2}}\right) e_{1d+1}\right] / \left[\Delta\tau_{d+\frac{1}{2}} \left(\Delta\tau_{d+\frac{1}{2}} + \Delta\tau_{d-\frac{1}{2}}\right)\right]
\end{aligned} \tag{5.9}$$

where

$$\begin{aligned}
e_{0d} &= 1 - \exp(-\Delta\tau_{d-\frac{1}{2}}) \\
e_{1d} &= \Delta\tau_{d-\frac{1}{2}} - 1 + \exp(-\Delta\tau_{d-\frac{1}{2}}) = \Delta\tau_{d-\frac{1}{2}} - e_{0d} \\
e_{2d} &= (\Delta\tau_{d-\frac{1}{2}})^2 - 2\Delta\tau_{d-\frac{1}{2}} + 2 - 2\exp(-\Delta\tau_{d-\frac{1}{2}}) = (\Delta\tau_{d-\frac{1}{2}})^2 - 2e_{1d}
\end{aligned} \tag{5.10}$$

For a linear source function the interpolation coefficients simplify to

$$\begin{aligned}
\alpha_d^- &= e_{0d} - e_{1d}/\Delta\tau_{d-\frac{1}{2}} \\
\beta_d^- &= e_{1d}/\Delta\tau_{d-\frac{1}{2}} \\
\gamma_d^- &= 0 \\
\alpha_d^+ &= 0
\end{aligned}$$

$$\begin{aligned} \beta_d^+ &= e_{1d+1}/\Delta\tau_{d+\frac{1}{2}} \\ \gamma_d^+ &= e_{0d+1} - e_{1d+1}/\Delta\tau_{d+\frac{1}{2}} \end{aligned} \tag{5.11}$$

Equations (5.4) and (5.5) may now be solved trivially for each angle and frequency. If one opts for second order accuracy, i.e. adopting the parabolic description eq. (5.7) for the source function, one needs to take some special care at the inner boundary for the incoming beam and at the outer boundary for the outgoing beam because α_D^- and β_D^- and β_0^+ and γ_0^+ are not defined (γ_D^- and α_0^+ are both zero). One way out is to assume first order accuracy (i.e. adopt a linear source function) for the above coefficients at $d = D$ and $d = 0$.

The above described technique is called *method of short characteristics* because it reduces the problem of solving the transfer equation to that of a series of analytical solutions over small intervals, where the source function behaves in a characteristic way. Note that the strong non-linear behaviour of the integrant $f(x) = S(x)e^{-x}$ (see eq. 5.1) is no longer a problem because of the analytical integration of the exponential term.

Though these lectures are not concerned with numerical techniques, we do show, for fun, how such a piece of code might look like. The listed routine is for a first order solution of the equation of transfer by means of the method of short characteristics and is programmed in IDL:

```
    pro short_characteristics_transfersolver, Src,tau,Int,Int_incident,ND
;
;   First order short characteristics transfer solver for outgoing beams
;   only, following Olson & Kunasz 1987, JQSRT 38, 325.
;
;   INPUT:
;
;   Src          -  source function grid
;   tau          -  optical depth grid where tau=0 at the observer
;   Int_incident -  incident specific intensity at the far side of the
;                   medium (as seen from the observer) for the outgoing
;                   beam. It should have the same units as Src
;   ND           -  nr of depth points; 0..ND-1 = outer..inner boundary
;
;   OUTPUT:
;
;   Int          -  specifice intensity grid for outgoing beam. The
;                   emerging intensity is Int(0)
;
    Nill = replicate(0.d0,ND)
    Int  = Nill & dInt = Nill & dtau_dph = Nill
;
    for d = 0,ND-2 do begin
       dtau_dph(d) = tau(d+1)-tau(d)
       e0_dp1      = 1.d0-exp(-dtau_dph(d))
       e1_dp1      = dtau_dph(d)-e0_dp1
       bdp         = e1_dp1/dtau_dph(d)
       cdp         = e0_dp1-e1_dp1/dtau_dph(d)
       dInt(d)     = bdp*Src(d)+cdp*Src(d+1)
    endfor
;
    Int(ND-1) = Int_incident
    for d = ND-2,0,-1 do begin
```

```
      Int(d)   = Int(d+1)*exp(-dtau_dph(d))+dInt(d)
    endfor
;
    return
    end
```

5.2 Feautrier method

The basis of Feautriers method is *i)* the use of separate transfer equations for inward and outward directed radiation, and *ii)* the introduction of a symmetric and an anti-symmetric average of the specific intensity. We first rewrite transfer equation eq. (4.36). We limit the beam angles μ to the half space $0 \leq \mu \leq 1$, and, for the moment, drop the standard notation of a subscript ν specifying the frequency dependence. We get

$$\pm\mu \frac{\mathrm{d}I(z, \pm\mu, \nu)}{\mathrm{d}\tau(z, \nu)} = I(z, \pm\mu, \nu) - S(z, \nu) \tag{5.12}$$

where $I(+\mu)$ and $I(-\mu)$ describe the outward and inward intensity. The symmetric and anti-symmetric average are defined as, respectively

$$u(z, \mu, \nu) = \frac{1}{2}\left[I(z, +\mu, \nu) + I(z, -\mu, \nu)\right] \tag{5.13}$$

and

$$v(z, \mu, \nu) = \frac{1}{2}\left[I(z, +\mu, \nu) - I(z, -\mu, \nu)\right] \tag{5.14}$$

Note that in a one-dimensional space, u would describe the mean intensity and v the flux. We may obtain a system of two 1^{st} order differential equations by adding the two equations (5.12), i.e.

$$\mu \frac{\mathrm{d}v}{\mathrm{d}\tau} = u - S \tag{5.15}$$

and subtracting them, i.e.

$$\mu \frac{\mathrm{d}u}{\mathrm{d}\tau} = v \tag{5.16}$$

Substitution of (5.16) in (5.15) yields the 2^{nd} order differential equation

$$\mu^2 \frac{\mathrm{d}^2 u}{\mathrm{d}\tau^2} = u - S \tag{5.17}$$

The boundary conditions are the incoming intensity $I(-\mu) = I_\mu^-$ at the edge $\tau = 0$ and the outgoing intensity $I(+\mu) = I_\mu^+$ at the edge $\tau = \tau_{\max}$. It has been found in many transfer problems that an accurate description of the boundary conditions is essential. The two boundary equations follow from (5.16) and are given by

$$\mu \left.\frac{\mathrm{d}u}{\mathrm{d}\tau}\right|_0 = u(0) - I_\mu^- \tag{5.18}$$

and

$$\mu \left.\frac{\mathrm{d}u}{\mathrm{d}\tau}\right|_{\tau_{\max}} = I_\mu^+ - u(\tau_{\max}) \tag{5.19}$$

— — —

To give an example of what the boundary conditions I_μ^- and I_μ^+ may look like, we use the model atmosphere as an example. In this case the incident intensity at the outer boundary $\tau = 0$ is $I_\mu^- \equiv 0$, such that

$$\mu \left.\frac{\mathrm{d}u}{\mathrm{d}\tau}\right|_0 = u(0) \tag{5.20}$$

The incoming intensity at the inner boundary $\tau = \tau_{\max}$ of the atmosphere follows from eq. (4.60) and $S_\nu = B_\nu$ (see § 6.2). This yields $I_\mu^+ = B_\nu(\tau_{\max}) + (\chi_\nu^{-1}|dB_\nu/dz|)_{\tau_{\max}}$. For $u(\tau_{\max}, \mu, \nu) = B_\nu(\tau_{\max})$ we then arrive at

$$\mu \left.\frac{\mathrm{d}u}{\mathrm{d}\tau}\right|_{\tau_{\max}} = \mu \left(\frac{1}{\chi_\nu}\left|\frac{\mathrm{d}B_\nu}{\mathrm{d}z}\right|\right)_{\tau_{\max}} \tag{5.21}$$

— — —

The next step is to formulate the system of differential equations (5.17), (5.18) and (5.19) as a set of difference equations using the *finite difference method*.

We choose a set of of optical depth points $\{\tau_d\}, d = 0, ..., D$ where $\tau_0 < \tau_1 < ... < \tau_D$; a set of angle points $\{\mu_m\}, m = 0, ..., M$ and a set of frequency points $\{\nu_n\}, n = 0, ..., N$. In this grid notation we may write the symmetric average of the intensity u as $u(z_d, \mu_m, \nu_n) = u_{dmn}$, and the source function as $S(z_d, \nu_n) = S_{dn}$. We apply the central difference approximation for arbitrary stepsize and find for the derivative

$$\left.\frac{\mathrm{d}u}{\mathrm{d}\tau}\right|_{d+\frac{1}{2}} = \frac{\Delta u_{d+\frac{1}{2}}}{\Delta\tau_{d+\frac{1}{2}}} = \frac{u_{d+1} - u_d}{\tau_{d+1} - \tau_d} \tag{5.22}$$

and for the second derivative

$$\left.\frac{\mathrm{d}^2u}{\mathrm{d}\tau^2}\right|_d = \frac{\left.\frac{\mathrm{d}u}{\mathrm{d}\tau}\right|_{d+\frac{1}{2}} - \left.\frac{\mathrm{d}u}{\mathrm{d}\tau}\right|_{d-\frac{1}{2}}}{\frac{1}{2}(\Delta\tau_{d+\frac{1}{2}} + \Delta\tau_{d-\frac{1}{2}})} \tag{5.23}$$

where we have introduced

$$\Delta\tau_{d\pm\frac{1}{2}} = \frac{1}{2}\left(\chi_{d\pm1} + \chi_d\right)|z_{d\pm1} - z_d| \tag{5.24}$$

and

$$\Delta\tau_d = \frac{1}{2}\left(\Delta\tau_{d-\frac{1}{2}} + \Delta\tau_{d+\frac{1}{2}}\right) \tag{5.25}$$

Equiped with these definitions we find for the difference representation of the 2nd order transfer equation (5.17)

$$\left(\frac{\mu_m^2}{\Delta\tau_{d-\frac{1}{2}}\Delta\tau_d}\right)u_{d-1} - \frac{\mu_m^2}{\Delta\tau_d}\left(\frac{1}{\Delta\tau_{d-\frac{1}{2}}} + \frac{1}{\Delta\tau_{d+\frac{1}{2}}}\right)u_d + \left(\frac{\mu_m^2}{\Delta\tau_d\Delta\tau_{d+\frac{1}{2}}}\right)u_{d+1} = u_d - S_d \tag{5.26}$$

This equation is valid at the depth points $d = 2, ..., D-1$. At the boundary $\tau = 0$ we have

$$\mu_m\frac{u_1 - u_0}{\Delta\tau_{\frac{1}{2}}} = u_0 - I_m^- \tag{5.27}$$

and at $\tau = \tau_{\max}$

$$\mu_m\frac{u_D - u_{D-1}}{\Delta\tau_{D-\frac{1}{2}}} = I_m^+ - u_D \tag{5.28}$$

Feautrier elimination scheme

The final step in the Feautrier method is to write the set difference equations (5.26), (5.27), and (5.28) as a matrix equation. To that end we introduce the vector $\mathbf{u}_d$ which for given beam angle μ and frequency ν has dimension D. The matrix equation can be expressed as

$$-\mathbf{A}_d\mathbf{u}_{d-1} + \mathbf{B}_d\mathbf{u}_d - \mathbf{C}_d\mathbf{u}_{d+1} = \mathbf{L}_d \tag{5.29}$$

where the $(D \times D)$ matrix $\mathbf{A}_d$ contains only *lower diagonal* terms; $\mathbf{B}_d$ only *diagonal* terms, and $\mathbf{C}_d$ only *upper diagonal* terms. Note that $\mathbf{A}_0 = 0$ and $\mathbf{C}_D = 0$. The right hand side vector $\mathbf{L}_d$ contains source terms. The structure of the system looks as follows

$$\begin{pmatrix} B_0 & -C_0 & & & & \\ -A_1 & B_1 & -C_1 & & & \\ & -A_2 & B_2 & -C_2 & & \\ & & ... & ... & ... & \\ & & & -A_{D-1} & B_{D-1} & -C_{D-1} \\ & & & & -A_D & B_D \end{pmatrix}\begin{pmatrix} u_0 \\ u_1 \\ u_2 \\ ... \\ u_{D-1} \\ u_D \end{pmatrix} = \begin{pmatrix} L_0 \\ L_1 \\ L_2 \\ ... \\ L_{D-1} \\ L_D \end{pmatrix} \tag{5.30}$$

The matrix has a tridiagonal structure. One may solve this type of matrix by an efficient forward-backward recursive sweep through the system. We start at $d = 0$, and express the symmetric average u_d in terms of u_{d+1}. If

$$u_d = D_d u_{d+1} + E_d \tag{5.31}$$

then

$$D_d = (B_d - A_d D_{d-1})^{-1} C_d \tag{5.32}$$

and

$$E_d = (B_d - A_d D_{d-1})^{-1}(L_d + A_d E_{d-1}) \tag{5.33}$$

On arrival at $d = D$, where $C_D = 0$, such that $D_D = 0$ we find that $u_D = E_D$. Now that u_D is known we may do a back substitution in (5.31) and recover all values for $\mathbf{u}_d$. If one is only interested in the emerging intensity at $\tau = 0$, at which in many transfer problems the incident intensity $I_\mu^- = 0$, one could start the recursive sweep at $d = D$ and one would not even require back substitution: $I_0^+ = 2u_0 - I_\mu^-$. A routine for solving tridiagonal systems is given in, for example, Numerical Recipes.

Exercise 5.1

a) Derive the set of difference equations (5.26), (5.27) and (5.28).

The boundary conditions (5.27) and (5.28) are only accurate to within first order. In this exercise we will derive the boundary conditions to within second order precission.

b) Show by means of a 'forward' Taylor series expansion to within second order that the boundary condition (5.27) at $\tau = 0$ can be improved to

$$\mu_m \frac{u_1 - u_0}{\Delta\tau_{\frac{1}{2}}} = u_0 - I_m^- + \frac{\Delta\tau_{\frac{1}{2}}}{2} \frac{(u_0 - S_0)}{\mu_m} \tag{5.34}$$

c) Show by means of a 'backward' Taylor series expansion to within second order that the boundary condition (5.28) at $\tau = \tau_{\max}$ can be improved to

$$\mu_m \frac{u_D - u_{D-1}}{\Delta\tau_{D-\frac{1}{2}}} = I_m^+ - u_D - \frac{\Delta\tau_{D-\frac{1}{2}}}{2} \frac{(u_D - S_D)}{\mu_m} \tag{5.35}$$

6

Radiation and matter

6.1 Introduction

The most important parameters describing the material medium are the mass density $\rho(\mathbf{r})$ and the temperature $T(\mathbf{r})$. The essential problem in the study of astrophysical media is to understand the interaction between the material medium and the radiation field. The basic quantities mentioned can be determined from the conservation laws of hydrodynamics. The density structure of the medium is determined by the conservation of mass and momentum; the temperature structure follows from the conservation of energy.

The description of the coupling between the gas (and/or dust) and the radiation field can be simplified considerably if the medium and the radiation are in some type of equilibrium.

Types of equilibrium between particles and photons

The most strict form of equilibrium is *thermodynamic equilibrium* (TE). A system that is in TE is in *thermal, mechanical, and chemical equilibrium*. So, there is no internal heat transport. All forces in the system are balanced. Finally, all chemical reactions are reversible with no net change in the amounts of reactants and products; its chemical composition does not change. In TE the medium is homogeneous and at rest. There are no gradients. Each process is in a microscopic equilibrium with its reverse process, i.e. all processes are in *detailed balance*. The radiation field is given by the Planck function, which only depends on T, and both the thermal velocity distribution of the particles and the distribution of particles over excitation and ionization states depend only on ρ and T.

The simple fact that we receive photons of the objects in which we are interested tells us that these media can not be in a state of TE. Because photons escape from these media it must be so that significant gradients are present in the quantities describing the medium. If we can not assume that the medium may be characterized as a whole with one value for ρ and one value for T, but if we are allowed to describe the state of the material medium *locally* using only

the value of $\rho(\mathbf{r})$ and $T(\mathbf{r})$, then we refer to the situation as being in *local thermodynamic equilibrium* (LTE). In LTE all atomic processes are still in detailed balance. However, the radiation field is *not* in equilibrium, but follows from the solution of the equation of transfer.

In those cases that even LTE is not valid, the medium is (per definition) in a state of *non local thermodynamic equilibrium* (NLTE). This is so when (at least) one microscopic process is not in detailed balance. The population of (at least) two excitation/ionization levels will deviate from their LTE values.

In this chapter we mainly focus on a description of TE and LTE.

6.2 Thermodynamic equilibrium

Thermodynamic equilibrium is realized in a cavity that is enclosed by isothermal walls, and that has relaxed to a situation of rest. As photons do not have a mass they can be absorbed and emitted by the walls of the cavity in arbitrary numbers. We expect that the number of photons and the distribution of these photons over frequency will reach an equilibrium state, characterized by the temperature T of the walls. The state of the radiation field (as well as that of the medium; but more about that later) in TE is therefore described by only one variable, the temperature T. This variable is independent of the nature of the walls or of the shape, texture, or size of the container.

The latter we may understand using a thought experiment. We connect the container to a second container of which the isothermal walls are also at the temperature T, and we carefully drill a hole in the containers such that their radiation fields are in contact. We place a filter that only allows radiation of frequency ν to pass through (see figure 6.1). If $I_\nu^1 \neq I_\nu^2$, then an energy flow would occur between both cavities. This would be in contradiction to the second law of thermodynamics (it would allow the possibility of a *perpetuum mobile*). From our experiment we may also conclude that the intensity in the enclosure is isotropic ($I_\nu \neq I_\nu(\mathbf{n})$) and homogeneous ($I_\nu \neq I_\nu(\mathbf{r})$). In other words

$$I_\nu = I_\nu(T) \equiv B_\nu(T) \tag{6.1}$$

In circa 1860 Gustav Robert Kirchhoff (1824–1887) measured the best approximation of this equilibrium radiation by making a small hole in the side of a closed box. He realized that radiation incident to the enclosure would be absorbed by the box, though perhaps only after very many reflections off the insides of the walls (provided the hole was sufficiently small). The box thus behaves as a perfect *blackbody*, i.e. a body that absorbs radiation of any frequency completely. Such a body is also the *best* possible emitter. The equilibrium radiation that emerges from the hole is therefore referred to as *blackbody radiation*. Max Planck (1858–1947) derived the function that describes the radiation field of a blackbody. We will discuss the shape of this function, after introducing Kirchhoff's law.

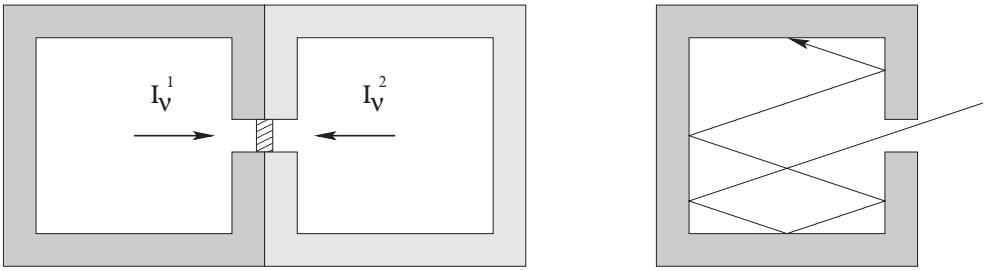


Figure 6.1: *Left panel: Two isothermal containers of which the walls have identical temperatures. A filter that only allows through radiation of frequency ν is in front of a small hole separating the enclosures. If the system is in equilibrium then the radiation fields in both containers must be identical. Right panel: Radiation incident to a small hole in a container will be absorbed by the walls, though perhaps only after many reflections. The radiation emitted by the insides of the walls (so also the radiation that emerges from the hole) will be an excellent approximation of* blackbody radiation.

Kirchhoff's law & the Kirchhoff-Planck relation

In a medium that is in TE there are no gradients, i.e. the specific intensity is homogeneous, isotropic and time independent. In that case eq. (4.13) implies that for all rays, for all frequencies, at all times

$$\eta_\nu = \chi_\nu I_\nu \tag{6.2}$$

This is *Kirchhoff's law*. Combining this law with eq. (6.1) results in

$$\eta_\nu = \chi_\nu B_\nu(T) \tag{6.3}$$

This is the *Kirchhoff-Planck relation*. So, the source function $S_\nu = \eta_\nu/\chi_\nu$ in TE is equal to the Planck function B_ν.

In summary: in TE I_ν= B_ν= S_ν. As the radiation field is isotropic and homogeneous, $I_\nu = J_\nu$ and $\mathcal{F}_\nu = 0$.

6.3 Planck function

The Planck function per frequency unit is given by

$$B_\nu(T) = \frac{2h\nu^3}{c^2}\frac{1}{e^{h\nu/kT}-1} \tag{6.4}$$

where h and k are Planck's constant and Boltzmann's constant respectively, and c is the speed of light. The dimensions of B_ν are erg cm^{-2} s^{-1} hz^{-1} sr^{-1}. A derivation of the Planck function is for instance given in Rybicki & Lightman. Using the relation

$$|B_\nu(T)\,d\nu| = |B_\lambda(T)\,d\lambda|\,, \tag{6.5}$$

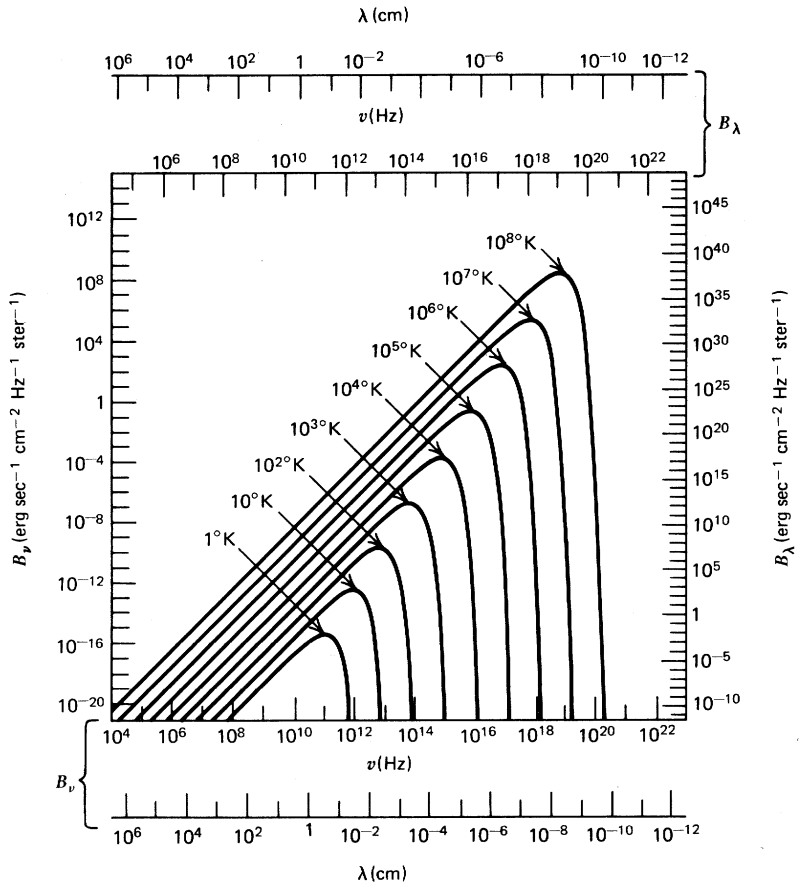


Figure 6.2: *The Planck curve B_ν for different temperatures. Note that the peak frequency (Wien's displacement law) and the integral of B_ν (Stefan-Boltzmann's law) increase with temperature.*

such that $B_\lambda = (c/\lambda^2)\, B_\nu$, we may rewrite the Planck function in wavelength units. We find

$$B_\lambda(T) = \frac{2hc^2}{\lambda^5} \frac{1}{e^{hc/\lambda kT} - 1} \tag{6.6}$$

of which the dimensions are erg cm^{-2} s^{-1} cm^{-1} sr^{-1}.

A series of Planck curves (in frequency unit) is given in figure 6.2. This reveals a number of striking properties of this function: *i)* The maximum of B_ν shifts toward higher frequencies if the temperature increases; *ii)* the Planck curves do not intersect: B_ν increases monotonous with temperature, therefore the frequency integrated or total Planck function will do so too, and *iii)* at frequencies much lower than the peak position $\partial \log B_\nu / \partial \log \nu = 2$.

Let us study the properties of the Planck function in more detail.

Wien's displacement law

Note that if we write $B_\nu(T) = T^3 q^3 \mathcal{B}(q)$, where $q = \nu/T$, that the function $\mathcal{B}(q)$ is a universal function. This proofs that the Planck curves will not intersect. The position of the maximum of the Planck curves follows from *Wien's displacement law*, and can be derived by taking the partial derivative $\partial B_\nu/\partial\nu = 0$. We find

$$h\nu_{\max} = 2.82144\, kT \qquad \text{or} \qquad \frac{\nu_{\max}}{T} = 5.87870 \times 10^{10} \quad \text{hz K}^{-1} \tag{6.7}$$

The peak of B_λ is at

$$\lambda_{\max} T = 0.28979 \quad \text{cm K} \tag{6.8}$$

Note: the maxima of B_ν and B_λ are *not* at the same position in the spectrum. For instance, the cosmic microwave background (CMB) temperature is 2.72 K (Planck Collaboration et al., 2016). It peaks at 160 GHz or 0.19 cm in frequency space and at 281 GHz or 0.11 cm in wavelength space.

Rayleigh-Jeans approximation

For frequencies low enough to have $h\nu/kT \ll 1$, the Planck function simplifies to the *Rayleigh-Jeans approximation*

$$B_\nu(T) \simeq \frac{2\nu^2 kT}{c^2} \qquad \text{or} \qquad B_\lambda(T) = \frac{2ckT}{\lambda^4} \tag{6.9}$$

This explains the linear parts at the left side of figure 6.2. The Rayleigh-Jeans approximation can almost always be applied in the radio regime. The coldest a natural system in space can be is about the CMB temperature (excluding black holes), implying that thermal emission of everything drops off very fast at frequencies much below ~ 100 GHz. The LOFAR radio-telescope observes in the 30−80 MHz and 120−240 MHz bands, i.e. far in the Rayleigh-Jeans regime of any thermal body in space. [1]

Wien's approximation

For frequencies high enough to have $h\nu/kT \gg 1$, the Planck function simplifies to *Wien's approximation*

$$B_\nu(T) \simeq \frac{2h\nu^3}{c^2} e^{-h\nu/kT} \qquad \text{or} \qquad B_\lambda(T) = \frac{2hc^2}{\lambda^5} e^{-hc/\lambda kT} \tag{6.10}$$

These are the steep parts at the right side of figure 6.2.

[1] Continuum emission at such low frequencies becomes dominated by particles which have been accelerated to very much above the thermal average by energetic processes. The most important such process for radio astronomy is synchrotron emission, caused by acceleration of electrons in a magnetic field (e.g. in jets, novae, and supernovae).

Stefan-Boltzmann's law

Integrating over the entire spectrum yields *Stefan-Boltzmann's law*

$$B \equiv \int_0^\infty B_\nu \, d\nu \equiv \int_0^\infty B_\lambda \, d\lambda = \frac{\sigma}{\pi} T^4 \tag{6.11}$$

where

$$\sigma = \frac{2\pi^5 k^4}{15 h^3 c^2} = 5.66961 \times 10^{-5} \quad \text{erg cm}^{-2}\,\text{K}^{-4}\,\text{s}^{-1} \tag{6.12}$$

is Stefan-Boltzmann's constant.

6.4 Laws describing the material medium in TE

Maxwellian velocity distribution

The probability, in TE, that a particle of mass m and temperature T has a velocity in the range $(\mathbf{v}, \mathbf{v} + \mathbf{dv})$ is given by the *Maxwellian velocity disbribution*

$$f(\mathbf{v})\, dv_x dv_y dv_z = \left(\frac{m}{2\pi kT}\right)^{3/2} \exp\left[-m(v_x^2 + v_y^2 + v_z^2)/2kT\right]\; dv_x dv_y dv_z \tag{6.13}$$

For each component the spread in velocities is thus given by a Gauss distribution. In the x direction for instance

$$f(v_x)\, dv_x = \frac{1}{\sqrt{\pi}} \left(\frac{m}{2kT}\right)^{1/2} \exp\left[-m v_x^2/2kT\right]\; dv_x \tag{6.14}$$

The most probable velocity in the x direction is $v_x = 0$. The root-mean-square velocity in this direction is

$$< v_x^2 >^{1/2} = \left[\frac{\int_0^\infty v_x^2 \, f(v_x)\, dv_x}{\int_0^\infty f(v_x)\, dv_x}\right]^{1/2} = \left(\frac{kT}{m}\right)^{1/2}. \tag{6.15}$$

To find the probability distribution in terms of speed v we must integrate over all direction components. We find

$$f(v)\, dv = \left(\frac{m}{2\pi kT}\right)^{3/2} \exp\left[-mv^2/2kT\right]\; 4\pi v^2\, dv \tag{6.16}$$

The dimensions of $f(v)$ are cm^{-1} sec and the integral of $f(v)$ is unity, similar to $f(\mathbf{v})$ and $f(v_x)$. It is not a Gauss distribution, but shows a 'tail' as a result of the v^2 term (see figure 6.3). The exponent of the distribution function is the ratio of the gas particle's kinetic

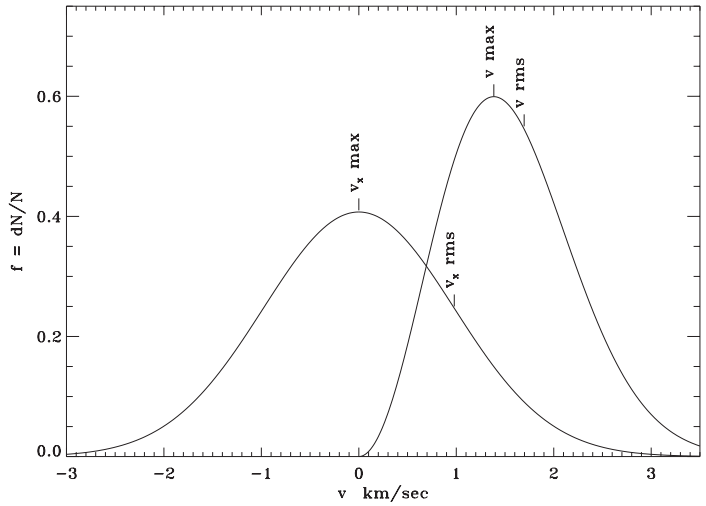


Figure 6.3: *The Maxwellian velocity distribution for iron at a temperature of 6 000 K, both in terms of one of the velocity components (for instance the one in the line of sight) and in terms of the absolute velocity v; $v_{\rm max}$ indicates the most probable speed, and $v_{\rm rms}$ the root-mean-square velocity.*

energy, $mv^2/2$, to the characteristic thermal energy, kT. It is relatively improbable that a significant number of the particles have an energy much greater or less than the thermal energy; the distribution peaks when these energies are equal, at a most probable speed

$$v = \left(\frac{2kT}{m}\right)^{1/2}. \tag{6.17}$$

The mean speed is

$$< v >= \left(\frac{8kT}{\pi m}\right)^{1/2} = 14.551 \left(\frac{T}{10^4 A}\right)^{1/2} \quad \mathrm{km\ s^{-1}}, \tag{6.18}$$

where A is the atomic weight (in amu) of the particle. The root-mean-square speed is

$$< v^2 >^{1/2}= \left(\frac{3kT}{m}\right)^{1/2} = 15.793 \left(\frac{T}{10^4 A}\right)^{1/2} \quad \mathrm{km\ s^{-1}}. \tag{6.19}$$

This implies an average thermal energy of the particle

$$\frac{1}{2}m\langle v^2\rangle = \frac{3}{2}kT. \tag{6.20}$$

The typical thermal energy of the particle is therefore $E_{\rm th} \sim kT$.

If the Maxwellian velocity distribution is valid one speaks of *kinetic equilibrium* (KE). In KE processes that exchange energy between particles (such as collisions between free particles)

occur rapid compared to the timescale on which processes occur that change kinetic conditions (such as collisional ionisations and recombinations). In almost all astrophysical media the free electrons are in KE. For this condition to be valid an electron that is ejected from an ion (for instance after a photoionization or a collisional ionization) needs to experience a large number of elastic collisions (usually with other electrons) before it suffers a non-elastic collision with an atom or ion. KE implies that the medium has a unique kinetic *electron temperature* $T_{\rm e}$. Atoms and ions also fulfill the Maxwell velocity distribution. Their matching kinetic temperature will be almost identical to $T_{\rm e}$ if the density is not too low and the temperature is not too high. Only in very rarefied plasmas such as the solar corona, where $T > 10^6$ K and $n_{\rm e} < 10^5$ cm^{-3}, substantial differences between the two may occur.

Boltzmann excitation equation

In TE the number density distribution of electronic transitions of atoms, ions, or molecules and of vibrational and rotational transitions of molecules over all discrete excitation states (the bound energy levels) is given by the *Boltzmann excitation equation*.

$$\frac{n_j}{n_i} = \frac{g_j}{g_i} \exp\left[-(E_j - E_i)/kT\right] = \frac{g_j}{g_i} \exp\left[-h\nu_{ij}/kT\right] \tag{6.21}$$

Here n_i is the number density per cm^3 in level i; g_i is the *statistical weight*, equivalent to the number of states with the same energy but different quantum numbers; E_i the *excitation energy* in erg, and $\nu_{ij} = \Delta E_{ij}/h$ the frequency in Hz corresponding to the energy difference $\Delta E_{ij} = E_j - E_i$.

It is customary that within each ionization stage the excitation energy is measured from the ground state up (see figure 6.4). The ionization energy is also measured per ionization stage from the ground state up. Usually one does not give the energy difference between levels in erg, but in electron volts (eV) or wavenumber (cm^{-1}). If we adopt the first unit one speaks of *excitation potential* . In this case E_j[erg] = 1.602192 10^{-12} E_j[eV] and ΔE_{ij}[eV] = 12 398.54 / λ_{ij}[Å]. If we adopt the last unit the wavelength of the transition can trivially be recovered using λ_{ij}[Å] = $10^8/(E_j$[cm^{-1}] - E_i[cm^{-1}]). Wavenumber simply provides the number of wavelengths that fit in a cm.

The Boltzmann equation can also be written as

$$\frac{n_i}{N} = \frac{g_i}{U(T)} \exp\left[-E_i/kT\right] \tag{6.22}$$

where $N = \sum_i n_i$ is the sum of populations over all levels, i.e. the total particle density per cm^3 of a given ionization stage, and $U(T)$ is the *partition function* of this ionization stage given by

$$U(T) \equiv \sum_i g_i \exp\left[-E_i/kT\right] \tag{6.23}$$

The above two equations are for instance used in traditional curve-of-growth analysis (see § 13.4).

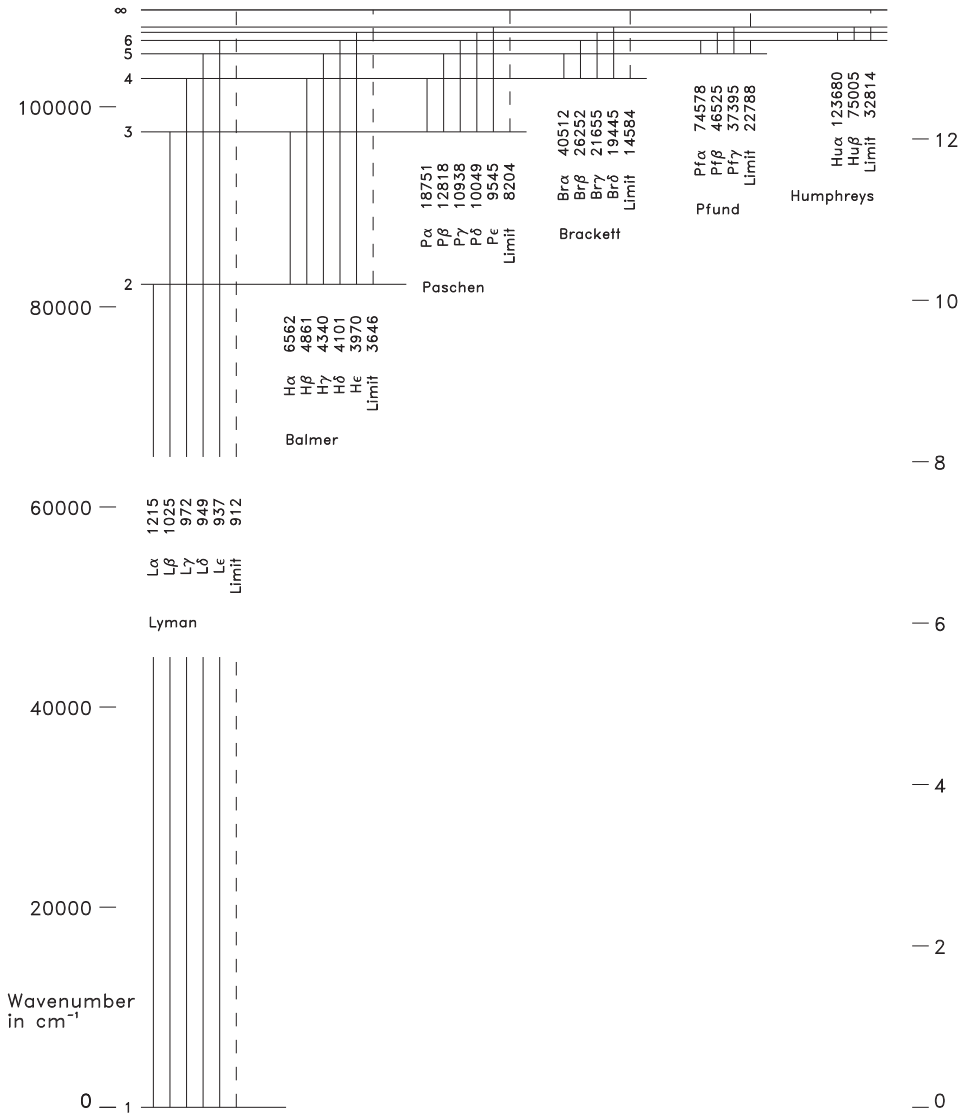


Figure 6.4: *Schematic picture of the energy level diagram or* Grotrian diagram *of hydrogen. The excitation energy of the levels is measured from the ground state up and is given in units of wavenumber on the left axis and in electron volts on the right axis. See table 7.1 for addiation information.*

Saha ionization equation

In TE the number density distribution over the ionization stages of an element is given by the *Saha ionization equation*

$$\begin{aligned} \frac{N_I}{N_{I+1}} &= n_{\rm e}\,\frac{U_I}{2U_{I+1}}\left(\frac{h^2}{2\pi m_{\rm e} kT}\right)^{3/2}\exp\left[E_I/kT\right] \\ &\equiv n_{\rm e}\tilde{\Phi}_I(T) \end{aligned} \tag{6.24}$$

where $n_{\rm e}$ is the electron density per cm^3 and $m_{\rm e}$ the electron mass; $N_I = \sum_i n_{i,I}$ is the particle density of ions in ionization stage I; U_I is the partition function of this ion; and E_I its ionization energy (in erg), i.e. the minimum energy required to liberate an electron from the ion. The factor 2 in front of the partition function U_{I+1} reflects the statistical weight of the liberated electron, that may have one out of two possible spin orientations.

If the temperature decreases at constant density (such that $n_{\rm e}$ remains more or less the same) the exponential term in the Saha ionization equation will make that N_I/N_{I+1} increases, i.e. the medium *recombines*. Essentially, the root-mean-square speed (Eq. 6.19) of the Maxwellian velocity distribution decreases, lowering the number of free electrons that can cause ionization. If the density decreases at constant temperature, $n_{\rm e}$ (more or less proportional to ρ) will decrease and therefore also N_I/N_{I+1}, i.e. the medium *ionizes*. Essentially, the lower electron density lowers the probability that a free electron is captured by an ion, while at the same time collisional ionizations are unimportant relative to photo-ionizations in liberating electrons from ions in the conditions typical for stellar photospheres (see section 15.2).

In many cases the particle density of the ground level $i = 0$ will dominate the particle density of the ion I, i.e. $N_I = \sum_i n_{i,I} \simeq n_{0,I}$. The Saha equation for ground levels only is given by

$$\frac{n_{0,I}}{n_{0,I+1}} = n_{\rm e}\frac{g_{0,I}}{2g_{0,I+1}}\left(\frac{h^2}{2\pi m_{\rm e} kT}\right)^{3/2}\exp\left[E_I/kT\right] \tag{6.25}$$

Finally, we may express the particle density n_i in terms of the particle density of the ground level of ion $I+1$ and the electron density by substitution of eq. (6.21) in (6.25). This gives

$$\begin{aligned} n_{i,I} &= n_{0,I+1} n_{\rm e}\frac{g_{i,I}}{2g_{0,I+1}}\left(\frac{h^2}{2\pi m_{\rm e} kT}\right)^{3/2}\exp\left[(E_I - E_i)/kT\right] \\ &\equiv n_{0,I+1} n_{\rm e}\,\Phi_{iI}(T) \end{aligned} \tag{6.26}$$

This is the *Saha-Boltzmann equation* or *equation of state* (EOS) of a gas in TE or LTE. In the latter case one should use the local values for temperature and density (see section 6.5). The equation can also be used to *define* LTE particle densities in a NLTE medium, i.e. using the values for $n_{0,I+1}$ and n_e that follow from imposing statistical equilibrium (see eq. 9.19). Deviations from LTE can thus be described by the *NLTE departure coefficient*

$$b_{i,I} \equiv \frac{n_{i,I}}{n_{i,I}^{\rm LTE}} \tag{6.27}$$

in which $n_{i,I}^{\rm LTE}$ is given by eq. (6.26). Henceforth we will denote LTE values of the level populations with a superscript $*$ (and not with LTE). However, we will do so only if this is required in the context of the discussion (if not, no superscript is used).

6.5 The TE equation of state for an ionized gas

Using the Saha-Boltzmann relation we may determine for each element the number densities of all ionization stages, and its contribution to the number of free electons in the plasma. Different elements will have widely different ionization energies; some will be neutral under given conditions, others will be singly or multiply ionized. For increasing temperature the transition from one to the next ionization stage is usually rather abrupt. This yields a sensitive diagnostic for the temperature (structure) of the medium, as it implies that the ratios in line strength of two successive ions (for instance He I & II, or Ca I & II) will rapidly change as a function of temperature (see § 2.1). Fundamental work in this field has been done at the start of the previous century by, among others, Saha, Anton Pannekoek , and Fowler & Milne (see figure 6.5).

In a normal stellar atmosphere hydrogen is by far the most abundant element, followed by helium ($N_{\rm He}/N_{\rm H} \sim 0.1$). The heavier elements have far smaller abundances (see table 16.2). In the solar atmosphere, which has a temperature of about 6 000 K, hydrogen is almost completely neutral. Free electrons mainly originate from "metals" such as Na, Mg, Al, Si, Ca, en Fe. At temperatures characteristic for A-type stars (10 000 K) hydrogen starts to ionize and becomes the dominant source of free electrons. At very high temperatures, typically for O- and B-type stars, also helium gets ionized and will contibute noticeably to the electron density.

In this section we will describe how the state of the gas in TE can be determined from the values of the temperature T and the total particle density N, or equivalently, from the total gas pressure $p_{\rm G}$ and the electron density $n_{\rm e}$.

Ideal gas law

As we will show in chapter 9 the total gas pressure is an important quantity in the description of a stellar atmosphere. If we assume that the gas is *ideal*, i.e. that there are no forces at work among the particles themselves – which is in almost all cases a valid assumption – then the gas pressure $p_{\rm G}$ and temperature T are related to the total particle number density N in the gas as (see also eq. 9.2)

$$p_{\rm G} = NkT \tag{6.28}$$

The total particle number density itself is the sum of all atoms, ions, and free electrons

$$N = N_{\rm atoms} + N_{\rm ions} + n_{\rm e} = N_N + n_{\rm e} \tag{6.29}$$

where N_N is the number density of nuclei.

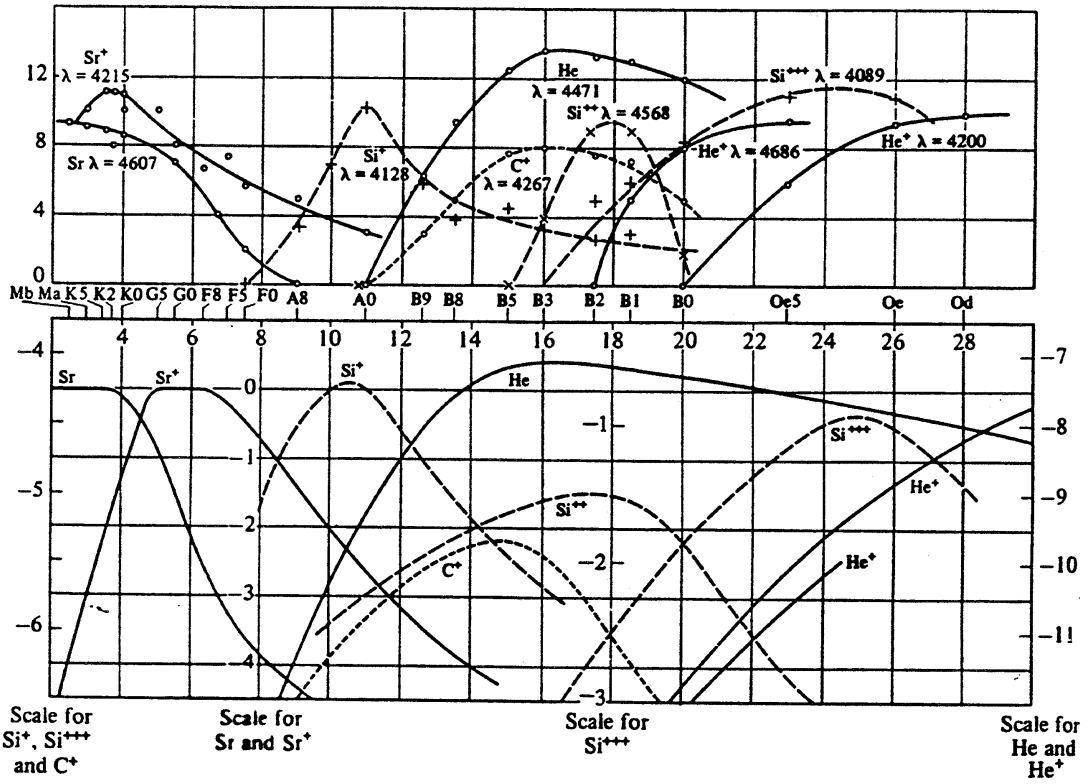


Figure 6.5: *The link between the spectral classification and the ionization of a gas in TE. Top: estimates (by eye, from photographic plates) of the strength of representative spectral lines as a function of the heuristically introduced spectral type. Bottom: ionization fractions for* $p_\mathrm{e} = n_\mathrm{e}\,k\,T = 131$ *dyne cm*$^{-2}$. *After Payne.*

Particle and charge conservation

If we define the *abundance* or *number abundance* A_k of a chemical element k as

$$N_k = A_k N_N \quad \text{where} \quad \sum_k A_k \equiv 1 \quad \text{such that} \quad \sum_k N_k = N_N \tag{6.30}$$

then either the last equality of the above equation, or

$$N_k = A_k(N - n_\mathrm{e}) \tag{6.31}$$

summarizes the constraint of *particle conservation*.

We will also require that the plasma is electrically neutral. In order for this to hold the total charge of the free electrons must be equal to the total ionic charge, i.e.

$$n_\mathrm{e} = \sum_k \sum_{j=1}^{J_k} j N_{jk} = \sum_k N_k \sum_{j=1}^{J_k} j q_{jk}(n_\mathrm{e}, T) = (N - n_\mathrm{e}) \sum_k A_k \sum_{j=1}^{J_k} j q_{jk}(n_\mathrm{e}, T) \tag{6.32}$$

where for each element we start the summation over ionization stages at $j = 1$, as the neutral atom $j = 0$ does not yield any free electrons, and end at the maximum ionization J_k. Note that we have changed the index label of the ionization stages from I to j. Equation (6.32) defines the condition of *charge conservation*. In this condition the *ionization fraction* $q_{jk} \equiv (N_{jk}/N_k)$ gives the fraction of atoms of element k in ionization stage j relative to the total number of atoms of species k.

In TE

$$\begin{aligned} q_{jk} &\equiv (N_{jk}/N_k)^* \\ &= \frac{(N_{jk}/N_{j+1,k})^* \cdots (N_{J-1,k}/N_{Jk})^*}{(N_{0k}/N_{1k})^* \cdots (N_{J-1,k}/N_{Jk})^* + (N_{J-2,k}/N_{J-1,k})^* (N_{J-1,k}/N_{Jk})^* + (N_{J-1,k}/N_{Jk})^* + 1} \\ &= \prod_{l=j}^{J_k-1} \left[n_{\rm e}\tilde{\Phi}_{lk}(T)\right] \Big/ \sum_{m=0}^{J_k} \prod_{l=m}^{J_k-1} \left[n_{\rm e}\tilde{\Phi}_{lk}(T)\right] \\ &\equiv P_{jk}(n_{\rm e}, T)/\, S_k(n_{\rm e}, T), \qquad (j = 1, \cdots, J_k) \end{aligned} \tag{6.33}$$

Note that the product term for $l = J_k$, which formally is not defined, is substituted by 1 in both the numerator and denominator. If we know $(n_{\rm e}, T)$, we may compute all q_{jk}'s and also, using eq. (6.32), N.

A pure hydrogen gas

For a pure hydrogen gas the system of equations (6.24), (6.31), and (6.32) may be solved analytically to obtain

$$n_{\rm e}(N, T) = \left[\left(N\tilde{\Phi}_{\rm H}(T) + 1\right)^{1/2} - 1\right] \Big/ \tilde{\Phi}_{\rm H}(T) \tag{6.34}$$

For an increasing density the ionization at constant temperature will decrease as the recombination rate increases as $N^+ n_{\rm e}$, while the ionization only scales with N^0 (see § 8.1).

General solution for the ionization of a gas in TE

In the general case the problem of determining $n_{\rm e}$ in a gas in TE, for given values of N and T requires an *iterative linearization procedure*. We will describe this procedure in fair detail because it is a simple example of the approach taken in more complicated problems in stellar atmospheres.

In a multi-dimensional linear perturbation or *Newton-Raphson* method we start with an initial estimate $\mathbf{x}^\circ = (x_1^\circ, \ldots, x_n^\circ)$. This is likely not the correct answer to our problem, therefore we assume that the true solution is given by a linear expansion $\mathbf{x} = \mathbf{x}^\circ + \delta\mathbf{x}^\circ$ where $\delta\mathbf{x}^\circ$ is to be determined in such a way as to satisfy exactly the equation at hand. In other words, we if

we seek the solution of the problem $f(\mathbf{x}) = 0$, we assume that

$$f(\mathbf{x}^\circ) + \sum_{i=1}^{n} \left.\frac{\partial f}{\partial x_i}\right|_{\mathbf{x}^\circ} \delta x_i^\circ = 0 \tag{6.35}$$

Because the equation is nonlinear, we cannot determine the exact value of this $\delta\mathbf{x}^\circ$, but need to start an iteration procedure in which we use $\mathbf{x}^\circ(\mathrm{new}) = \mathbf{x}^\circ(\mathrm{old}) + \delta\mathbf{x}^\circ$ to update the value of $\delta\mathbf{x}^\circ$. The convergence of this procedure is quadratic (if our original estimate lies within the range of convergence) and one can obtain the result to within the desired accuracy quickly.

The only equation we need to solve by linearization is eq. (6.32), where the ionization fractions $q_{jk}(n_\mathrm{e}, T)$ are given by (6.33). First order perturbation of both sides gives

$$\begin{aligned} n_\mathrm{e} + \delta n_\mathrm{e} &\approx (N - n_\mathrm{e}^\circ)\tilde{\Sigma}(n_\mathrm{e}^\circ, T) + \tilde{\Sigma}(n_\mathrm{e}^\circ, T) \left.\frac{\partial(N - n_\mathrm{e})}{\partial n_\mathrm{e}}\right|_{n_\mathrm{e}^\circ} \delta n_\mathrm{e} + (N - n_\mathrm{e}^\circ) \left.\frac{\partial\tilde{\Sigma}}{\partial n_\mathrm{e}}\right|_{n_\mathrm{e}^\circ} \delta n_\mathrm{e} \\ &= (N - n_\mathrm{e}^\circ - \delta n_\mathrm{e})\tilde{\Sigma}(n_\mathrm{e}^\circ, T) + (N - n_\mathrm{e}^\circ) \left.\frac{\partial\tilde{\Sigma}}{\partial n_\mathrm{e}}\right|_{n_\mathrm{e}^\circ} \delta n_\mathrm{e} \end{aligned} \tag{6.36}$$

such that

$$\delta n_\mathrm{e} \approx \left[(N - n_\mathrm{e}^\circ)\tilde{\Sigma} - n_\mathrm{e}^\circ\right] \left[1 + \tilde{\Sigma} - (N - n_\mathrm{e}^\circ)\left(\partial\tilde{\Sigma}/\partial n_\mathrm{e}\right)\right]^{-1} \tag{6.37}$$

where

$$\tilde{\Sigma}(n_\mathrm{e}^\circ, T) \equiv \sum_k A_k S_k^{-1}(n_\mathrm{e}^\circ, T) \sum_{j=1}^{J_k} j P_{jk}(n_\mathrm{e}^\circ, T) \tag{6.38}$$

The derivative $\partial\tilde{\Sigma}/\partial n_\mathrm{e}$ can be obtained analytically

$$\frac{\partial\tilde{\Sigma}}{\partial n_\mathrm{e}} = \sum_k A_k \left[S_k^{-1} \sum_j^{J_k} j \frac{\partial P_{jk}}{\partial n_\mathrm{e}} - S_k^{-2} \frac{\partial S_k}{\partial n_\mathrm{e}} \sum_j^{J_k} j P_{jk}\right] \tag{6.39}$$

where $(\partial P_{jk}/\partial n_\mathrm{e})$ and $(\partial S_k/\partial n_\mathrm{e})$ are easily determined if one considers that

$$P_{jk}(n_\mathrm{e}, T) \equiv \prod_{l=j}^{J_k - 1} \left[n_\mathrm{e}\tilde{\Phi}_{lk}(T)\right] = n_\mathrm{e}^{(J_k - j)} \Pi_{jk}(T) \tag{6.40}$$

and

$$S_k(n_\mathrm{e}, T) \equiv \sum_{j=0}^{J_k} P_{jk}(n_\mathrm{e}, T) = \sum_{j=0}^{J_k} n_\mathrm{e}^{(J_k - j)} \Pi_{jk}(T) \tag{6.41}$$

The derivatives that occur in a linearization procedure can always be determined numerically. However, in quite some cases they can also be derived analytically. The latter is the preferred way to go as experience teaches that this way one has more "control" over the computation.

— — —

Finally, after one has obtained $n_{\rm e}$ using the above procedure, and, as a by-product all ionization fractions q_{jk}, one may calculate all desired number particle densities from (for instance) Boltzmann equation (6.22). From $N_{jk} = q_{jk} A_k (N - n_{\rm e})$ one gets

$$n_{ijk} = \frac{g_{ijk}}{U_{jk}(T)} \exp[-E_{ijk}/kT]\, q_{jk}(n_{\rm e}, T) A_k (N - n_{\rm e}) \tag{6.42}$$

This completes the calculation of the TE equation of state for an ionized gas.

6.6 Temperature definitions

Based on the above discussions we may define six characteristic temperatures that are often used in astrophysics. The first two are related to the Planck function; the next two to the equations of Boltzmann and Saha. One is related to the total flux emitted at the surface of the radiating body (this is $T_{\rm eff}$) and one to the power from a source received by a telescope.

Brightness temperature

The *brightness temperature* $T_{\rm b}$ is the temperature for which the Planck function, at the frequency at which is measured, reproduces the observed specific intensity

$$I_\nu^{\rm obs} = B_\nu(T_{\rm b}) \tag{6.43}$$

The brightness temperature is often used in radio astronomy, where in most cases the Rayleigh-Jeans approximation is valid such that

$$T_{\rm b} \equiv \frac{h\nu/k}{\ln\left[1 + 2h\nu^3/c^2 I_\nu^{\rm obs}\right]} \simeq \frac{c^2}{2\nu^2 k} I_\nu^{\rm obs} \tag{6.44}$$

In general $T_{\rm b}$ will depend on frequency. Only if the source emits a blackbody spectrum $T_{\rm b}$ will be the same for all frequencies.

If we consider a medium at a constant temperature T that emits according to the Planck function, the solution of the equation of transfer (vgl. 4.49) in the Rayleigh-Jeans limit is

$$T_{\rm b}(0) = T_{\rm b}(\tau_\nu) e^{-\tau_\nu} + T\left[1 - e^{-\tau_\nu}\right] \qquad \forall\ \frac{h\nu}{kT} \ll 1 \tag{6.45}$$

For large optical depth in the medium the brightness temperature will be equal to the temperature of the material, i.e. $T_{\rm b}(0) = T$.

Color temperature

The *color temperature* T_c is the temperature for which the Planck function, at the frequency $\nu_\circ$ at which is measured, reproduces the slope of the observed spectrum

$$\left.\frac{\mathrm{d}I_\nu^{\mathrm{obs}}}{\mathrm{d}\nu}\right|_{\nu_\circ} = \left.\frac{\mathrm{d}B_\nu(T_c)}{\mathrm{d}\nu}\right|_{\nu_\circ} \tag{6.46}$$

Often the shape of a (part of the) observed stellar spectrum more or less resembles that of a Planck curve. However stars are not spatially resolved (save for the sun and a few nearby asymptotic giant stars and supergiants) such that only the flux can be observed. Therefore, flux is what is commonly used. In everyday life a number of apparent magnitudes m is measured, as discussed in § 2.6, from which colors are composed. The $B{-}V$ color, for instance, follows from

$$B{-}V = -2.5\log\left[\frac{\int_0^\infty \mathcal{F}_\nu(d)\,\mathcal{S}_B(\nu)\,d\nu}{\int_0^\infty \mathcal{F}_\nu(d)\,\mathcal{S}_V(\nu)\,d\nu}\right] + 2.5\,(C_B - C_V) \tag{6.47}$$

where $\mathcal{F}_\nu(d)$ is the stellar flux at the location of the earth, and $\mathcal{S}_m$ has the meaning as given in § 2.6. In figure 6.6 the $U{-}B$ vs. $B{-}V$ diagram for the MK spectral types is repeated (see also figure 2.11). Also given is the result of equation 6.47 assuming a Planck curve for the radiation field (solid line). Note that the $U{-}B$ vs. $B{-}V$ dependence is almost linear, and that it does not agree very well with the relation given by the MK calibration. The reason for the difference, obviously, is that a real star emits a spectrum that deviates from the Planck curve. Especially the Balmer jump has a strong influence on the U magnitude, therefore on $U{-}B$.

For a simple estimate of the values of the color indices using the Planck function we assume that the UBV respons functions can be replaced by δ functions at 3600, 4400, and 5500 Å, respectively. Using the constants C_U, C_B, and C_V as given in table 2.2 (Landelt-Börnstein 1982) we get for the colors of the Planck-curves

$$B{-}V = -2.5\log\left[\frac{B_B}{B_V}\right] + 0.22 \qquad \text{en} \qquad U{-}B = -2.5\log\left[\frac{B_U}{B_B}\right] - 1.04 \tag{6.48}$$

For this result figure 6.6 also gives the relation between $U{-}B$ and $B{-}V$ (dashed line). Because now we did not convolve with the respons functions $\mathcal{S}_m$ the relation differs from the solid line.

Color temperatures are often used to characterize stars and gaseous planets. For relations between colors and effective temperature derived from detailed atmosphere models, see the tables in the appendix.

Excitation temperature

The excitation temperature is determined by comparison of observed level populations of an atom, ion, or molecule, using the Boltzmann equation (eq 6.21). Obviously, the measured

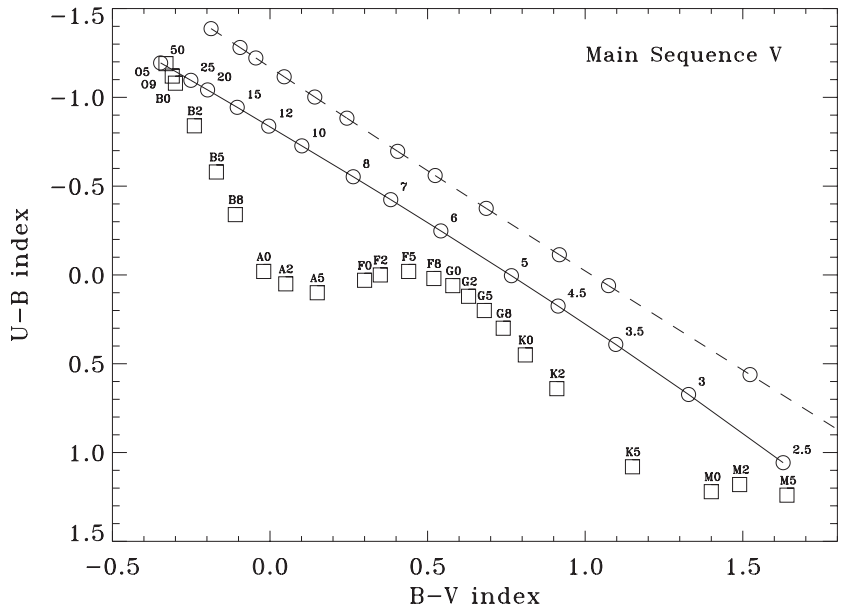


Figure 6.6: *Relation between the color indices $U-B$ and $B-V$, assuming the flux distribution of the star is given by the Planck function (solid line). If one does not convolve with the proper response functions, but adopts a peaked profile at the centre wavelength of the filter, one obtains monochromatic values for the color indices (dashed line).*

populations need not be in LTE. To conserve a description of the measured population ratios in terms of the Boltzmann equation, we introduce the excitation temperature $T_{\rm ext}$. This temperature is given by

$$\left(\frac{n_j}{n_i}\right)^{\rm obs} = \frac{b_j}{b_i}\frac{n_j^*}{n_i^*} = \frac{b_j}{b_i}\frac{g_j}{g_i}\exp\left[-h\nu_{ij}/kT\right] \equiv \frac{g_j}{g_i}\exp\left[-h\nu_{ij}/k\,T_{\rm ext}\right] \tag{6.49}$$

If $b_u/b_l = 1$, such as is the case of TE and LTE, then $T = T_{\rm ext}$. The excitation temperature can be determined from measurements of the equivalent width of two spectral lines (see e.g. § 13.4). The derived value of the excitation temperature can be rather ambiguous. Even for different line pairs of the same ion the excitation temperatures may differ significantly. Sometimes this is due to uncertainties in the oscillator strength (see § 7.3). More often it is because lines may form in different regions of the atmosphere, where temperatures and/or departures from LTE may be different. The excitation temperatures derived from molecular lines (in cool stellar atmospheres) are usually lower than those found from lines of atomic or ionized gas. This is so because molecular lines typically originate in higher atmospheric layers, where T is relatively low.

Ionization temperature

In analogy to the excitation temperature, the *ionization temperature* is the temperature for which an observed ionization ratio fulfills the Saha equation (eq. 6.24)

$$\left(\frac{N_I}{N_{I+1}}\right)^{\text{obs}} \equiv n_{\text{e}} \frac{U_I}{2U_{I+1}} \left(\frac{h^2}{2\pi m_{\text{e}} k\, T_{\text{ion}}}\right)^{3/2} \exp\left[E_I/k\, T_{\text{ion}}\right] = n_{\text{e}} \tilde{\Phi}_I(\, T_{\text{ion}}) \tag{6.50}$$

This temperature can only be determined if the electron density n_{e} is known. In LTE it holds that $T = T_{\text{ion}}$.

Effective temperature

The *effective temperature* T_{eff} of a source is a measure of the total flux emitted by that source.

$$\mathcal{F}^{+}_{\text{source}} = \sigma\, T^4_{\text{eff}} \tag{6.51}$$

So, it is the temperature that an isotropic black surface should have such that the total outward directed flux $\mathcal{F}^{+} = \pi B = \pi \int_0^\infty B_\nu \, d\nu$, per cm^2, is equal to the flux emitted by the source, again per cm^2.

Antenna temperature

Sub-millimeter and radio astronomers often use the concept of an antenna temperature. The motions of charged particles in a resistor in an electrical circuit cause the resistor to generate noise and to reach a certain temperature. The frequency spectrum of the power of this noise only depends on the temperature of the resistor (and not, for instance, on the material of which the resistor is made). So, by equating the power the antenna is receiving to the temperature of a resistor that is connected to the antenna by power lines one can measure the antenna temperature.

We assume the effective surface of the antenna to be A_{eff} cm^2. The definition of the flux (see section 3.4) tells us that we receive a power $P_\nu = \mathcal{F}_\nu\, A_{\text{eff}}$ from the source. Radio receivers, for instance a dipole antenna, are typically sensitive to only one polarization direction. Therefore, only part of the total power incident to the effective surface is measured. For an unpolarized electromagnetic wave $E(\omega) = E_\circ e^{i\omega t}$ the flux scales with the square of the strength of the electrical field. For the mean power of the polarization component $E(\omega) = E_\circ \cos\omega$ it then follows that:

$$\langle P_\nu \rangle \propto \left\langle \cos^2 \omega \right\rangle = \frac{1}{2\pi} \int_0^{2\pi} \cos^2 \omega \, d\omega = \frac{1}{2\pi} \left[\frac{\omega}{2} + \frac{1}{4} \sin 2\omega\right]_0^{2\pi} = \frac{1}{2} \tag{6.52}$$

The power absorbed by the antenna is thus given by $P_\nu = 1/2\, A_{\text{eff}}\, \mathcal{F}_\nu$.

For the power that flows through the transmission wire, as well as the noise that is generated in the resistor, one finds analogues to Planck's law[2]

$$P_\nu = \frac{h\nu}{e^{h\nu/kT} - 1} \tag{6.53}$$

In the Rayleigh-Jeans limit it follows that $P_\nu = kT$. This is *Nyquist's law*, stating that a resistor at temperature T produces a noise of which the power per unit bandwidth is kT. Equating this to the power that is received from the source leads to an antenna temperature

$$T_\mathrm{A} = \frac{\mathcal{F}_\nu \, A_\mathrm{eff}}{2k}. \tag{6.54}$$

This result connects the flux of radiation from the source to an antenna temperature using the effective surface of the telescope.

To establish a relation between the antenna temperature and the physically more relevant brightness temperature we use relation 3.29, where we assume that the solid angle $d\omega$ observed of (part of) the source emits an isotropic radiation field I_ν. This leads to

$$T_\mathrm{A} = \frac{A_\mathrm{eff}}{\lambda^2} \, T_\mathrm{b} \, d\omega. \tag{6.55}$$

6.7 Approximations in describing the interaction of matter and radiation

Local thermodynamic equilibrium (LTE)

In a medium in LTE the material medium can be described using the equations of Maxwell, Boltzmann and Saha, using the local values of $\rho(\mathbf{r})$ en $T(\mathbf{r})$. However, the radiation field is not in equilibrium, i.e. it may deviate from the Planck function. LTE is an often applied generalization of TE (see chapter 11).

In LTE the assumption of detailed balance in all microscopic processes involving atomic transitions is still valid (see chapters 7 and 8). An important consequence is that the source function that is ascribed to each such process is still given by the Kirchhoff-Planck relation (eq. 6.3), i.e.

$$S_\nu(\mathbf{r}) = B_\nu(T(\mathbf{r})) \tag{6.56}$$

An example of a source function for which this relation holds is the *line source function* S_ν^ℓ that describes the microscopic transition between two bound energy levels (see § 7.3). Also the source functions of bound-free and free-free continuum processes (see chapter 8) are given by Kirchhoff-Planck. In these continuum processes energy is either added to or taken away from

[2] This result follows from a one-dimensional consideration of what leads to the Planck function in three dimensions.

the local thermal energy pool of the medium. These source functions are therefore referred to as *thermal source functions*. A process in LTE that can not be described by eq. (6.56) is that of *scattering*. This is to be expected as the mean intensity $J_\nu(\mathbf{r})$ – the relevant radiation quantity if we assume that the scattering process is isotropic – can deviate from the local emitted thermal radiation field $B_\nu(T(\mathbf{r}))$.

In summary: in LTE

$$n_{ijk} = n_{ijk}(\rho(\mathbf{r}), T(\mathbf{r})) \tag{6.57}$$

and

$$S_\nu^\ell(\mathbf{r}) = B_\nu(T(\mathbf{r})) \qquad I_\nu(\mathbf{r}, \mathbf{n}) \neq B_\nu(T(\mathbf{r})) \qquad J_\nu(\mathbf{r}) \neq B_\nu(T(\mathbf{r})) \qquad \mathcal{F}_\nu(\mathbf{r}) \neq 0 \tag{6.58}$$

The first equality implies that also *complete redistribution* applies (see §§ 7.1 and 7.3).

The essence of LTE is that of the two types of processes that determine the state of the material medium, i.e. the collisionally controlled "true" absorption and emission processes on the one hand and the scattering processes on the other hand, it is the collisional processes that dominate. In other words: the mean-free-path of photons for absorption $\ell_\nu = 1/\kappa_\nu$ is smaller than the mean-free-path for scattering $\ell_\nu = 1/\sigma_\nu$ (see eq. 4.28). According to the assumptions that define LTE the state of the material medium is determined by such a small TE-volume $V(\mathbf{r})$ that the fact that the temperature and density elsewhere in the medium may be different is not noticed by the excitation and ionization state of the gas in $V(\mathbf{r})$. Still, the radiation that is present in this volume does contain information about far away regions. However, the photons in the radiation field that have been created elsewhere and have reached $\mathbf{r}$ directly or through scattering do not bring enough information about these far away places to push the number populations of the energy levels out of equilibrium.

Non-local thermodynamic equilibrium (NLTE)

NLTE is per definition a state that deviates from LTE. In NLTE the state of the particles at position $\mathbf{r}$ is no longer controlled only by the local values of density and temperature but will also depend on the radiation field $J_\nu(\mathbf{r})$, i.e.

$$n_{ijk} = n_{ijk}(\rho(\mathbf{r}), T(\mathbf{r}), J_\nu(\mathbf{r})) \tag{6.59}$$

Departures from LTE typically occur if

- the particle densities are low, such that collisional processes will no longer dominate over scattering processes and photon conversion processes (see § 4.1), and, *in addition*
- the radiation field strongly deviates from the local Planck function, i.e. $J_\nu(\mathbf{r}) \neq B_\nu(T(\mathbf{r}))$.

The exact assumptions that enter in NLTE are not formally defined. Usually it is implied that *statistical equilibrium* is valid (see § 9.2); that the Maxwell velocity distribution holds, and

that complete redistribution applies. The level occupation numbers may deviate from the local Saha-Boltzmann values. The extinction coefficient may differ from the LTE value and the line source function may differ from the Planck function.

Exercise 6.1

Why is a blackbody the best possible emitter?

Exercise 6.2

a) Give the net flux $\mathcal{F}_\nu$ in a medium that is in TE and in which the specific intensity is I_ν.

b) Do we observe objects in the sky that are in TE?

Exercise 6.3

a) Give the derivation of eq. (6.6) using eq. (6.4), and the relation (6.5)

b) Show that $B_\nu \downarrow 0$ for $T \downarrow 0$, and that $B_\nu \uparrow \infty$ for $T \uparrow \infty \ \forall \, \nu$.

c) Derive eq. (6.11) and the constant (6.12), given that

$$\int_0^\infty \frac{x^3}{e^x - 1}\, dx = \frac{\pi^4}{15} \tag{6.60}$$

Exercise 6.4

a) The number of photons per cm^3 per Hz in a Planck radiation field is $n(\nu) = u_\nu / h\nu$, where u_ν (see eq. 3.15) is the energy density of the Planck radiation field. Show that the total number of photons per cm^3 in the Planck radiation field

$$n_{\rm tot} = \frac{16\pi\zeta(3)k^3}{c^3h^3}\, T^3 \simeq 20\, T^3 \tag{6.61}$$

given that

$$\int_0^\infty \frac{x^2}{e^x - 1}\, dx = 2\,\zeta(3) \tag{6.62}$$

where $\zeta(3) = 1.202057$ is Apéry's constant.

b) Compute the mean energy per photon in a Planck radiation field.

Exercise 6.5

a) Derive eq. (6.16) from (6.13).

b) Show that both distributions are normalized. Remember that

$$\int_{-\infty}^{+\infty} e^{-x^2/a^2}\, dx = a\sqrt{\pi} \qquad \text{and} \qquad \int_{0}^{+\infty} x^2 e^{-x^2/a^2}\, dx = a^3\sqrt{\pi}/4 \tag{6.63}$$

c) Show that the mean particle energy is given by

$$< \frac{1}{2}mv^2 >= \frac{3}{2}kT \tag{6.64}$$

if we give that

$$\int_{0}^{+\infty} x^4 e^{-x^2/a^2}\, dx = \frac{3}{8}a^5\sqrt{\pi} \tag{6.65}$$

d) Show that the most probable velocity of eq. (6.14) is given by $v_x = 0$ and of eq. (6.16) by $v = (2kT/m)^{1/2}$.

Exercise 6.6

Show that the ideal gas law eq. (6.28) is equivalent to (see also eq. 9.2)

$$p = \frac{\mathcal{R}\rho T}{\mu}, \tag{6.66}$$

where ρ is the density, $\mathcal{R} \equiv k/m_{\rm amu} = 8.314 \times 10^7$ erg K^{-1} mol^{-1} is the gas constant, and μ is the mean molecular weight of the gas in atomic mass units $m_{\rm amu}$ (so, the mean mass per free particle in $m_{\rm amu}$).

Exercise 6.7

a) Show that for a pure hydrogen gas in TE the analytical solution of the electron density in terms of $N_{\rm H}$ and T is given by

$$n_{\rm e}(N_{\rm H}, T) = \left[\left(4N_{\rm H}\tilde{\Phi}_{\rm H}(T) + 1\right)^{1/2} - 1\right] \Big/ 2\tilde{\Phi}_{\rm H}(T), \tag{6.67}$$

where $N_{\rm H} = N_N = N^0 + N^+ = N - n_{\rm e}$.

b) Show that in the general case, i.e. for any atomic gas in TE, the perturbation term δn_e in terms of N_N and T is given by

$$\delta n_\mathrm{e} = \left[N_N \tilde{\Sigma} - n_\mathrm{e}^\circ\right] \left[1 - N_N \left(\partial\tilde{\Sigma}/\partial n_\mathrm{e}\right)\right]^{-1}. \tag{6.68}$$

Exercise 6.8

The Time of Decoupling in the early universe refers to the era of recombination of the hot ionized gas in the expanding young cosmos. Before decoupling, Thomson scattering by free electrons caused a high opacity in the medium keeping the electrons and photons in thermal equilibrium. After decoupling, the loss of free electrons caused a drop in opacity, freeing the photons to roam unhindered throughout a newly transparent universe – hence the term 'decoupling'. As the majority of baryonic gas is in the form of hydrogen, recombination usually refers to the recombination of hydrogen gas. The primordial mass fraction of hydrogen is $X_\mathrm{prim} = 0.75$. For simplicity, we adopt that only hydrogen supplies free electrons for the cosmic gas (i.e. helium is neutral).

The temperature at the time of recombination can be estimated through the use of the Saha equation (6.24) for neutral and ionized hydrogen, where we use $U_1 \equiv U_\mathrm{I} = 2$ and $U_2 \equiv U_\mathrm{II} = 1$. The ionisation energy of hydrogen is E_I = 13.6 eV. We define q to be the fraction of hydrogen atoms that are ionised, i.e.

$$q = \frac{N_\mathrm{II}}{N_\mathrm{I} + N_\mathrm{II}}, \tag{6.69}$$

hence

$$\frac{N_\mathrm{II}}{N_\mathrm{I}} = \frac{q}{1-q}, \tag{6.70}$$

where N_I and N_II are the density of neutral and ionized hydrogen. In the present-day universe the baryonic density is $n_{\mathrm{b},\circ} = \rho_{\mathrm{b},\circ}/\mu\, m_\mathrm{H} = 5.4 \times 10^{-6}\,\mathrm{cm}^{-3}$, where the mean atomic weight is $\mu = 1.23$, and the temperature of the cosmic background radiation $T_\circ = 2.725$ K. We define the scale factor in the present-day universe to be $R_\circ = 1$. At the time of recombination, the universe was already matter-dominated, implying $n_{\mathrm{b},\circ} = R^3\, n_\mathrm{b}$, where n_b is the density at the time when the scale factor was R. The temperature at the time when the scale factor was R is given by $T = T_\circ/R$. Recall that the cosmic redshift is given by

$$z = \frac{1}{R} - 1. \tag{6.71}$$

a) Why might one consider the term 'recombination' in this context as oddly inappropriate?

b) Express the baryonic density n_b in the hydrogen density $n_\mathrm{H} = n_{\mathrm{H_I}} + n_{\mathrm{H_{II}}}$, assuming X is constant throughout cosmic time.

c) Use Saha equation (6.24) to derive an expression that links the scale factor R to the ionization fraction q, for the known present-day quantities $n_{\mathrm{b},\circ}$ and $T_\circ$.

d) Let us assume that the time interval of decoupling spans the era where q drops from 0.9 to 0.1 and that the surface of last scattering, from which the cosmic microwave background photons arriving at Earth were last scattered, corresponds to $q = 0.1$. Use for instance Excel to compute the scale factor at these two q values, hence the redshifts and temperatures at the start and end of recombination. (It may be convenient to introduce $R' = 10^3\, R$ for your Excel calculation, in which case the solutions are in the range $0.5 \leq R' \leq 1.0$).

To let you compare your findings to the results of the *Wilkinson Microwave Anisotropy Probe*: WMAP finds for the redshift at the time of decoupling $z_{\rm dec} = 1089 \pm 1$, and a temperature at the end of recombination of $T = T_{\rm o}\,(1 + z_{\rm dec})$ = 2970 K.

Exercise 6.9

a) Does the brightness temperature of a radio source depend on its distance?

b) Can one measure the brightness temperature of a point source (i.e. an object that is not spatially resolved) such as a star? Can one measure $T_{\rm b}$ for an extended source, such as a nebula, if this source is not in TE?

Exercise 6.10

What requirements have to be fulfilled in order to obtain the stellar temperature from two-color photometry? For which layer of the star would this temperature be representative?

Exercise 6.11

The displacement law of Wien (eq. 6.7) can also be used to derive the temperature of a star. Which characteristic temperature does this give: $T_{\rm b}$, $T_{\rm c}$ or $T_{\rm eff}$?

Exercise 6.12

Use the table to fill in – with a simple yes or no – whether for each of the three equilibria (TE, LTE, or NLTE) it holds that

a) the flux is equal to zero;

Equilibrium	$F_\nu = 0$	$S_\nu^{tot} = B_\nu$	$S_\nu^{th} = B_\nu$	$S_\nu^{line} = B_\nu$	Maxwell	Saha Boltzmann
TE						
LTE						
NLTE						

b) the total (i.e. continuum + line) source function is equal to the Planck function B_ν;

c) the thermal source function is equal to B_ν;

d) the line source functions are equal to B_ν;

e) the velocity distribution of the free elections is equal to the Maxwell distribution

f) the excitation and ionization equilibria are given by the equations of Boltzmann and Saha.

7

Discrete processes

So far we have given a macroscopic description of emission and extinction processes. In this chapter we focus on a microscopic specification. Between two energy levels five different processes may occur

1. *spontaneous radiative de-excitations*
   requires: nothing
2. *radiative excitations*
   requires: photon with the correct energy
3. *induced radiative de-excitations*
   requires: photon with the correct energy
4. *collisional excitations*
   requires: passing particle with sufficient energy
5. *collisional de-excitations*
   requires: passing particle

These processes take place between bound-bound (or bb) transitions as well as between bound-free (or bf) and free-free (or ff) transitions. More general, they occur in any system in which an exchange between internal energy and radiation is possible. So in all systems in which energy states can be defined, either discretely or continuously distributed over energy.

In this chapter we will examine these five processes for the bound-bound transitions between discrete energy levels. There are several types of such discrete states. We will mainly discuss the energy levels in the electron configuration of atoms and ions. Examples of other systems that have energy levels are *i)* shared electron configurations of molecules; *ii)* rotation of atoms in molecules; *iii)* vibration of atoms in molecules; *iv)* vibrations in the latice structure of solid state particles; and *v)* hadron configurations of atomic nuclei. The nature of these configurations and the selection rules that apply for them, and that follow from Pauli's exclusion principle for fermions, will not be treated here.

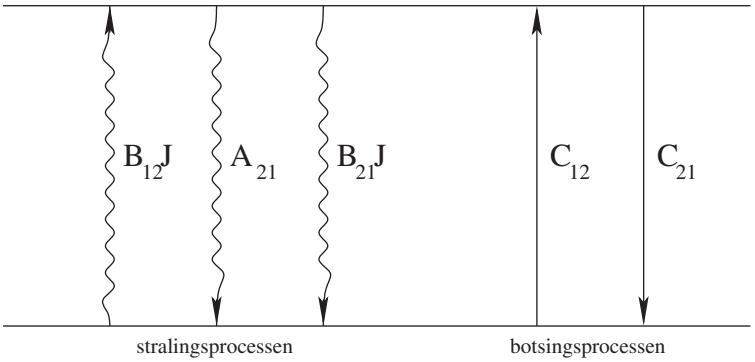


Figure 7.1: *Schematic overview of the five possible bound-bound transitions between two energy levels.*

The listed processes affect the number density of particles in a given state n_i. Let P_{ij} be the total rate of transition from state i to state j by all processes, then (in a static medium, see also section 9.2),

$$\frac{\partial n_i}{\partial t} = -n_i \, P_{ij} \tag{7.1}$$

is the number of particles per cm^3 per second that transition from state i to state j, hence P_{ij} has unit s^{-1}. Likewise,

$$\frac{\partial n_i}{\partial t} = n_j \, P_{ji} \tag{7.2}$$

is the number of particles per cm^3 per second that transition from state j to our state i. We will now specify these rates P_{ij} (and P_{ji}) in detail for each of the possible processes.

7.1 From a macroscopic toward a microscopic description

Spontaneous radiative de-excitation

In a *spontaneous radiative de-excitation* or *spontaneous emission* a particle in an excited energy state u ($upper$) returns spontaneously to a lower energy state l ($lower$) by emitting a photon. The rate at which this process occurs is defined by the *Einstein coefficient for spontaneous emission* A_{ul} (dimension sec^{-1}). It is an atomic parameter and does not depend on the properties of the medium in which the atom or ion is embedded (such as pressure, temperature, or radiation field), but is determined solely by the nature of the transition.

The number of spontaneous de-excitations in cm^{-3} sec^{-1} is given by

$$n_u \, R_{ul} = n_u \, A_{ul} \int \psi_\nu \, d\nu = n_u \, A_{ul} \tag{7.3}$$

where n_u is the number density of particles in level u in cm^{-3} and ψ_ν the normalized profile function for spontaneous emission (see below for some discussion on the profile function). The dimension of ψ_ν is hz^{-1}. R_{ul} is the rate for spontaneous de-excitation in sec^{-1}.

Radiative excitation

In a *radiative excitation* or *photoexcitation* a particle in energy state l is excited to a higher energy state u after absorption of a photon. The rate for this process depends on atomic properties, denoted by B_{lu}, as well as on the number of available photons of suitable frequency.

Due to some lack of sharpness of the energy levels, a certain spread in the energy that is required to induce the transition is present. We therefore introduce the *profile function* for excitation ϕ_ν of which the dimension is hz^{-1}, that describes this allowed spread in frequency (see § 13.2). This function is sharply peaked around the central frequency $\nu_{lu} = (E_u - E_l)/h$. The profile function, similar to ψ_ν, is normalized such that

$$\int_0^\infty \phi(\nu)\, d\nu \equiv 1 \tag{7.4}$$

The number of radiative excitations in cm^{-3} sec^{-1} is given by

$$n_l\, R_{lu} = n_l\, B_{lu} \int \phi_\nu\, J_\nu\, d\nu \equiv n_l\, B_{lu}\, \bar{J}_{lu} \tag{7.5}$$

where n_l is the number density of particles in level l in cm^{-3}; B_{lu} is the *Einstein coefficient for extinction* in erg^{-1} cm^2 hz, and $\bar{J}_{lu}$ is the integral of the mean intensity over the profile function of the transition. Note that the mean intensity is here the relevant radiation quantity: the particles are only interested in the total number of suitable photons, and not in the direction from which these photons arrive.

Induced radiative de-excitation

For Einstein to be able to derive the radiation formula of Planck he introduced a third radiative process, that of *induced de-excitation* or *stimulated emission*. In this process an incident radiation field triggers an emission, such that a particle in an excited state u returns to a lower energy state l.

The number of induced radiative de-excitations in cm^{-3} sec^{-1} is given by

$$n_u\, R_{ul} = n_u\, B_{ul} \int \varphi_\nu\, J_\nu\, d\nu \equiv n_u\, B_{ul}\, \bar{J}_{ul} \tag{7.6}$$

where B_{ul} is the *Einstein coefficient for stimulated emission*, and φ_ν is the normalized profile function for induced de-excitation. The photons that are emitted in induced radiative de-excitations have the same direction as do the incident photons that trigger the process.

In TE ϕ_ν, ψ_ν and φ_ν are identical as in that case detailed balance is valid, which applies for each frequency separately. In less strict circumstances this need not be the case. If, however, there is *no* correlation between the frequency of the incident photon (the one that causes the excitation) and that of the emitted photon (the one that is emitted in the de-excitation process), then $\phi_\nu = \varphi_\nu\,(= \psi_\nu)$ therefore $\bar{J}_{lu} = \bar{J}_{ul}$. This situation is known as *complete redistribution* and it is justified if the atoms/ions in a gas are so strongly perturbed by collisions during the excitation/emission process that the excited electrons are randomly redistributed over the substates of the upper level of the transition.

Collisional excitation

In a gas (consisting of atoms, ions, and electrons) of sufficient density a variety of collisions will occur, which may cause excitations and ionization. As soon as the gas is somewhat ionized, collisions of ions with charged particles will dominate over collisions with neutral particles. This is because the Coulomb interaction (governing the interaction of charged particles) reaches much farther than does the Van der Waals interaction (governing the interaction of neutral particles). As a rule electron-ion collisions are much more efficient than are ion-ion collisions. The reason is that the collision rate is proportional to the relative velocity of the colliding particles. In kinetic equilibrium the root-mean-square velocity is inversely proportional to the root of the mass of the particle (see eq. 6.19). Electrons thus move $(m_{\rm H}A/m_{\rm e})^{1/2} \sim 43A^{1/2}$ times as fast as do ions of atomic weight A.

Even in those parts of the solar atmosphere where the ionization fraction of hydrogen is only $\sim 10^{-5}$ one may assume that the relevant collisional processes are dominated by collisions with free electrons. This is partly so due to contributions of elements that have low ionization energies, such as iron and magnesium. Only in particularly cool environments, such as the atmospheres of extremely cool stars (late M-type stars such as AGB stars), brown dwarfs or gas-rich exoplanets, in proto-planetary disks and molecular clouds, collisions with neutral hydrogen atoms or molecular hydrogen may play a role, or even dominate.

The number of *collisional excitations* in cm^{-3} sec^{-1} is given by

$$n_l\, C_{lu} = n_l\, n_{\rm e} \int_{v_\circ}^{\infty} \sigma_{lu}(v)\, f(v)\, v\, dv \equiv n_l\, n_{\rm e}\, q_{lu}(T) \tag{7.7}$$

where $v_\circ$ is the threshold velocity (providing the minimum required kinetic energy $E_\circ = \frac{1}{2}mv_\circ^2 = h\nu_{lu}$), $n_{\rm e}$ the electron density per cm^3, and $\sigma_{lu}(v)$ the electron collision cross-section of the process in cm^2. In kinetic equilibrium, the velocity distribution $f(v)$ is that according to Maxwell (eq. 6.16). The collision cross-sections are for the most part poorly known. They are determined in experiments or by means of quantum mechanical calculations. We give an often used formula based on a dipole approximation, deduced by Van Regemorter

$$q_{lu}(T) \approx p_1 \left(\frac{E_0}{kT}\right)^{p_2} T^{-3/2}\, e^{-E_0/kT}\, f_{lu} \tag{7.8}$$

where $(p_1, p_2) = (2.16, -1.68)$ for atoms and $(3.9, -1)$ for ions (see e.g. Jefferies). The result is proportional to the oscillator strength f_{lu} of the line transition (see § 7.3), and only applies to allowed transitions. The rates for forbidden transitions (which have very small values of f_{lu}) are usually very poorly known, but most certainly are not as low as implied by Van Regemorters formula. The term $T^{-3/2}\exp(-E_\circ/kT)$ finds its origin in the Maxwellian velocity distribution (eq. 6.16); the exponential term dominates the rate, such that it increases rapidly with temperature. For a fixed temperature, the rate increases for decreasing E_0 as a lower threshold for the kinetic energy implies that a larger fraction of the free electrons (distributed according to Maxwell) can participate. This is very relevant for transitions between high lying energy levels, i.e. transitions close to the continuum (see figure 6.4). These have only small level-to-level energies $E_0 = h\nu_{lu}$; which typically correspond to line wavelengths in the infrared or submillimetre. These levels are strongly coupled by collisions.

Collisional de-excitation

The rate of *collisional de-excitations* can be obtained directly from a detailed balance argument. In KE the occupation numbers of the upper and lower level are related according to

$$n_l^*\, C_{lu} = n_u^*\, C_{ul} \tag{7.9}$$

In this equilibrium, set by collisions and only dependent on temperature, the ratio $(n_u/n_l)^*$ fulfills the Boltzmann equation (6.21).

The number of collisional de-excitations in $\mathrm{cm}^{-3}\,\mathrm{sec}^{-1}$ is then given by

$$n_u\, C_{ul} = n_u \left(\frac{n_l}{n_u}\right)^* C_{lu} = n_u\, n_\mathrm{e}\, q_{lu}(T)\, \frac{g_l}{g_u}\, e^{+E_\circ/kT} \tag{7.10}$$

Note that the rate for collisional de-excitations is much less sensitive to the temperature than is that of collisional excitations (the exponential term has dropped out). Roughly one could say that each encounter with an electron suffices to accelerate it, by transferring excitation energy.

Maxwellian averaged collision strength

We provide an alternative expression for the rate of collisional excitations using the *Maxwellian averaged collision strength* Υ, also referred to as the effective collisional strength. It holds that

$$q_{lu}(T) = \frac{8.629 \times 10^{-6}}{g_l\, T^{1/2}}\, e^{-E_{lu}/kT}\, \Upsilon_{lu}(T). \tag{7.11}$$

For the rate of collisional de-excitations it follows that

$$q_{ul}(T) = \frac{g_l}{g_u} e^{+E_{lu}/kT}\, q_{lu}(T) = \frac{8.629 \times 10^{-6}}{g_u\, T^{1/2}}\, \Upsilon_{lu}(T). \tag{7.12}$$

7.2 Einstein relations

The Einstein coefficients A_{ul}, B_{lu} and B_{ul} can be expressed in each other, assuming that thermodynamic equilibrium is valid.

In TE each specific process is in detailed balance: for each two energy levels the number of radiative excitations equals the number of radiative de-excitations

$$n_l\, B_{lu}\, \bar{J}_{lu} = n_u\, A_{ul} + n_u\, B_{ul}\, \bar{J}_{lu} \tag{7.13}$$

If we isolate the profile integrated mean intensity $\bar{J}_{lu}$, and use that in TE the occupation of the levels is given by Boltzmann (eq. 6.21), then

$$\begin{aligned} \bar{J}_{lu} &= \frac{A_{ul}/B_{ul}}{(n_l/n_u)\,(B_{lu}/B_{ul}) - 1} \\ &= \frac{A_{ul}/B_{ul}}{(g_l B_{lu}/g_u B_{ul})\, e^{h\nu_{lu}/kT} - 1} \end{aligned} \tag{7.14}$$

In TE $J_\nu = B_\nu$. As B_ν is a function that changes very little over the sharply peaked profile function ϕ_ν, we may also assume that $\bar{J}_{lu} = B_{\nu_{lu}}$. To fulfill this equation at each temperature, it must hold that

$$A_{ul} = \frac{2h\nu_{lu}^3}{c^2}\, B_{ul} \qquad \text{and} \qquad g_l\, B_{lu} = g_u\, B_{ul} \tag{7.15}$$

These are the *Einstein relations*. They connect atomic properties, i.e. properties that are independent of external circumstances. We therefore must conclude that they are universally valid, i.e. also in media that are not in TE, or for cases in which $\bar{J}_{lu} \neq B_{\nu_{lu}}$.

7.3 Relation between Einstein coefficients and η_ν^ℓ and χ_ν^ℓ

The emission coefficient η_ν defines the emitted energy per cm^3 per second per hz per solid angle (see eq. 4.7). We may derive this coefficient for spontaneous emission processes in a simple way. The energy emitted in all directions per cm^3 per second at all frequencies relevant for the transition is $h\nu_{lu}\, n_u A_{ul}$. To get the emitted energy per hz we need to multiply with the profile function for spontaneous emission ψ_ν (which has dimension hz^{-1}), and, finally, to get the emission per unit solid angle we need to divide by the total solid angle $\Omega = 4\pi$. This yields for the *line emission coefficient*

$$\eta_\nu^\ell = \frac{h\nu_{lu}}{4\pi}\, n_u\, A_{ul}\, \psi_\nu \tag{7.16}$$

Let us now derive an expression for the extinction coefficient χ_ν. The total energy that is absorbed in a volume element dV in a time interval dt by radiative excitation is

$$\begin{aligned} dE_\nu^{\rm tot} &= 4\pi \cdot h\nu_{lu}\, n_l\, B_{lu}\, \bar{J}_{lu}\, dV\, dt \\ &= 4\pi \cdot h\nu_{lu}\, n_l\, B_{lu} \left[\int \left\{ \frac{1}{4\pi} \oint I_\nu\, d\omega \right\} \phi_\nu\, d\nu \right] dV\, dt \end{aligned} \tag{7.17}$$

where the 4π is the integration over all directions and the last equality at the right-hand side follows from eq. (7.6) and (3.6). So, the energy dE_ν that is absorbed in a volume $dV = dOds$ in a bandwidth $d\nu$ in a time interval dt, from an incident beam of opening angle $d\omega$ and specific intensity I_ν is

$$dE_\nu = \frac{h\nu_{lu}}{4\pi}\, n_l\, B_{lu}\, \phi_\nu\, I_\nu\, dO\, ds\, d\omega\, d\nu\, dt \tag{7.18}$$

Using the definition of the extinction coefficient χ_ν (eq. 4.3) we find

$$\chi_\nu^{\rm excitation} = \frac{h\nu_{lu}}{4\pi}\, n_l\, B_{lu}\, \phi_\nu \tag{7.19}$$

Finally, let us turn to stimulated emissions. Usually this process is not added to that of spontaneous emission, but is treated in terms of a "negative extinction". The reason is that induced excitation is much more similar to radiative excitation than it is to spontaneous de-excitation. First because – if complete redistribution is valid – it is proportional to $\bar{J}_{lu}$. Second, the emission is in the same direction as that of the incident radiation by which it is induced (therefore the term "negative extinction"). In fact, radiative excitation and stimulated emission always occur together. We find for the *line extinction coefficient*

$$\chi_\nu^\ell = \frac{h\nu_{lu}}{4\pi}\, [\, n_l\, B_{lu}\, \phi_\nu - n_u\, B_{ul}\, \varphi_\nu\,] \tag{7.20}$$

Oscillator strength

To quantify the line emission and extinction coefficient we only need one of the Einstein coefficients (which needs to be determined by either measurement or calculation). The others follow from the Einstein relations eq. (7.15). Usually A_{ul} is tabulated. There is a fourth parameter that is used to quantify line strength: the *oscillator strength* f_{lu}.

This parameter is used in the classic description of a spectral line as a damped harmonic oscillator, according to which the extinction coefficient (or cross-section) per particle $\alpha_{lu}(\nu)$ (dimensions cm^2) is given by

$$\alpha_{lu}(\nu) = \frac{\pi e^2}{m_{\rm e} c}\, f_{lu}\, \phi_\nu \tag{7.21}$$

where $(\pi e^2/m_{\rm e}c)$ = 0.02654 cm^2 hz and ϕ_ν is the *Lorentz profile* or *damping profile* of which the dimension is hz^{-1}. (see § 13.2). The oscillator strength does not feature in the classical theory. A quantum mechanical treatment shows that the transition probabilities for different

transitions may differ by orders of magnitude. The dimensionless oscillator strength is thus introduced as a correction on the classical value.

In the above description the extinction coefficient is given by

$$\chi_\nu^{\rm excitation} = \alpha_{lu}(\nu)\, n_l = \frac{\pi e^2}{m_{\rm e} c}\, f_{lu}\, n_l\, \phi_\nu \tag{7.22}$$

The oscillator strength is related to the Einstein coefficient B_{lu} through the total extinction coefficient (or cross-section) per particle

$$\alpha_{lu} \equiv \int_0^\infty \alpha_{lu}(\nu)\, d\nu = \frac{\pi e^2}{m_{\rm e} c}\, f_{lu} = \frac{h\nu_{lu}}{4\pi}\, B_{lu} \tag{7.23}$$

of which the dimensions are (note!) $\mathrm{cm}^2\,\mathrm{hz}$. For resonance lines, such as Lyα, a classical oscillator is a good approximation and $f_{lu} \sim 1$. Other observed allowed lines in spectra typically have $10^{-4} \lesssim f_{lu} \lesssim 10^{-1}$. Forbidden transitions have $f_{lu} \lesssim 10^{-6}$. The oscillator strengths of a number of important hydrogen lines are given in table 7.1.

Finally: usually also $g_l f_{lu}$ is tabulated, known as the *gf-value*. In the literature, this quantity is often (erroneously) referred to as the "transition probability".

Kramers formula

For hydrogen an analytical expression for f_{lu} has been derived by Kramers

$$f_{\rm K}(l,u) = \frac{32}{3\pi\sqrt{3}} \left(\frac{1}{l^2} - \frac{1}{u^2}\right)^{-3} \left(\frac{1}{l^5\, u^3}\right) \tag{7.24}$$

where l and u are the principal quantum numbers of the lower and upper level respectively. The *Kramer formula* is a semi-classical result. A full quantum mechanical derivation results in a correction factor, the bound-bound Gaunt factor $g_{\rm I}$, that is of order unity. The oscillator strength is then given by

$$f(l,u) = g_{\rm I}(l,u)\, f_{\rm K}(l,u) \tag{7.25}$$

For the lines of a given spectral series, for instance the Lyman series ($1 \to u$) or Balmer series ($2 \to u$), the α-line ($l \to l+1$) will have the largest oscillator strength, followed by the β-line ($l \to l+2$), *etcetera*. This is reflected in the line strengths, e.g. Hα is almost always stronger than is Hβ, which in turn is stronger than Hγ, *etcetera*.

For hydrogenic ions

$$f_{lu} = Z^4\, f_{lu}(Z=1) \tag{7.26}$$

Table 7.1: *Wavelength λ_{lu}, oscillator strength f_{lu}, and Einstein coefficient A_{ul} for important transitions in hydrogen. See figure 6.4 for the Grotrian diagram*

Wavelength in Å

Series	l	u=2	3	4	5	6	7	8
Ly	1	1215.67	1025.72	972.54	949.74	937.80	930.75	926.23
H	2		6562.80	4861.32	4340.46	4101.73	3970.07	3889.05
Pa	3			18751.0	12818.1	10938.1	10049.4	9545.98
Br	4				40512.0	26252.0	21655.0	19445.6
Pf	5					74578	46525	37395
Hu	6						123680	75005

Oscillator strength

Series	l	u=2	3	4	5	6	7	8
Ly	1	0.4162	0.07910	0.02899	0.01394	0.007799	0.004814	0.003183
H	2		0.6407	0.1193	0.04467	0.02209	0.01270	0.008036
Pa	3			0.8421	0.1506	0.05584	0.02768	0.01604
Br	4				1.038	0.1793	0.06549	0.03230
Pf	5					1.231	0.2069	0.07448
Hu	6						1.424	0.2340

Einsteincoefficient A_{ul} in sec^{-1}

Series	l	u=2	3	4	5	6	7	8
Ly	1	$4.699\,10^8$	$5.575\,10^7$	$1.278\,10^7$	$4.125\,10^6$	$1.644\,10^6$	$7.568\,10^5$	$3.869\,10^5$
H	2		$4.410\,10^7$	$8.419\,10^6$	$2.530\,10^6$	$9.732\,10^5$	$4.389\,10^5$	$2.215\,10^5$
Pa	3			$8.986\,10^6$	$2.201\,10^6$	$7.783\,10^5$	$3.358\,10^5$	$1.651\,10^5$
Br	4				$2.699\,10^6$	$7.711\,10^5$	$3.041\,10^5$	$1.424\,10^5$
Pf	5					$1.025\,10^6$	$3.253\,10^5$	$1.388\,10^5$
Hu	6						$4.561\,10^5$	$1.561\,10^5$

Line source function

The *line source function*, using eq. (7.16) and (7.20), is given by

$$S_\nu^\ell = \frac{\eta_\nu^\ell}{\chi_\nu^\ell} = \frac{n_u\, A_{ul}\, \psi_\nu}{n_l\, B_{lu}\, \phi_\nu - n_u\, B_{ul}\, \varphi_\nu} \tag{7.27}$$

As the Einstein relations do also apply for conditions other than TE, the general form of the line source function (and in fact also of the source function of any arbitrary radiative transition) is

$$S_\nu^\ell = \frac{2h\nu_{lu}^3}{c^2} \frac{1}{(g_u\, n_l)/(g_l\, n_u) - 1} \tag{7.28}$$

where we have assumed complete redistribution, i.e. $\phi_\nu = \varphi_\nu = \psi_\nu$. The frequency dependence of the line source function is in effect extremely weak as the factor ν^3 varies only slowly compared with the sharply peaked profile function ϕ_ν. S_ν^ℓ is often referred to as a *frequency-independent* source function, i.e. $S_\nu^\ell = S^\ell$. In TE we may apply Boltzmann's law such that $S_\nu^\ell = B_\nu$. This recovers the Kirchhoff-Planck relation (eq. 6.3).

We end this section with two useful relations. The ratio of the number of spontaneous and induced de-excitations in TE is

$$\frac{n_u\, A_{ul}}{n_u\, B_{ul}\, \bar{J}_{lu}} = \frac{A_{ul}}{B_{ul}\, \bar{J}_{lu}} = \frac{2h\nu^3}{c^2 B_\nu} = e^{h\nu/kT} - 1 \tag{7.29}$$

So, if $h\nu/kT \gg 1$ (Wien's approximation) the majority of de-excitations will be spontaneous, and if $h\nu/kT \ll 1$ (Rayleigh-Jeans's approximation) the dominant downward radiative process is induced emission – provided that the populations and radiation field are of the order of what is expected in TE.

The ratio of the number of spontaneous de-excitations and the total radiative de-excitation rate in TE is

$$\frac{n_u\, A_{ul}}{n_u\, A_{ul} + n_u\, B_{ul}\, \bar{J}_{lu}} = \frac{A_{ul}}{A_{ul} + B_{ul}\, \bar{J}_{lu}} = 1 - e^{-h\nu/kT} \tag{7.30}$$

This term is referred to as *the correction factor for stimulated emission*. In Wien's approximation ($h\nu/kT \gg 1$), downward radiative transitions in TE or LTE are dominated by spontaneous de-excitations. In the Rayleigh-Jeans approximation ($h\nu/kT \ll 1$) induced de-excitations dominate as $B_{ul}/(A_{ul} + B_{ul}\, \bar{J}_{lu}) = \exp(-h\nu/kT) \simeq 1 - h\nu/kT \simeq 1$.

Exercise 7.1

a) At time $t = 0$ the population of an excited level is $n_u(0)$. How does the population of this level decrease over time if *only* spontaneous emission to a lower level l occurs and the Einstein coefficient of this process is A_{ul}?

b) If de-excitations are possible to multiple lower levels the transition probabilities should be added. Show that if $\Gamma_u \equiv \sum_{i<u} A_{ui}$ the *lifetime* (in sec) of the particle in state u is given by

$$< t >= \frac{1}{\Gamma_u} \tag{7.31}$$

Exercise 7.2

Use Kramers' formula eq. 7.24 to show that the α-line of hydrogen is the strongest absorption line of a series (i.e. of all lines with fixed lower level).

Exercise 7.3

a) Show that in NLTE, assuming complete redistribution, the line extinction coefficient per particle corrected for stimulated emissions is given by

$$\alpha_{lu}^{\rm cse}(\nu) = \frac{\pi e^2}{m_{\rm e} c} f_{lu} \left[1 - \frac{n_u \, g_l}{n_l \, g_u} \right] \phi_\nu \tag{7.32}$$

b) Show that in TE and LTE the line extinction coefficient is given by

$$\alpha_{lu}^{\rm cse}(\nu) = \frac{\pi e^2}{m_{\rm e} c} f_{lu} \left[1 - e^{-h\nu_{lu}/kT} \right] \phi_\nu \tag{7.33}$$

Exercise 7.4

This could be a nice exam question. The fine structure lines of neutral oxygen arise from the fact that the ground electronic state splits up into three fine structure levels. The lines are forbidden. Relevant atomic data for the lines are provided in table 7.2.

a) Make a sketch of the energy level structure of the oxygen ground electronic state with the three line transitions and calculate their wavelength.

Designation i	E_i [K]	g_i
$^3P_2$	0.0	5.0
$^3P_1$	227.7	3.0
$^3P_0$	326.6	1.0

Designation $i \to j$	A_{ij} [s^{-1}]	q_{ij} [$cm^3\ s^{-1}$]
$^3P_1 - {}^3P_2$	8.9×10^{-5}	1.5×10^{-10}
$^3P_0 - {}^3P_2$	1.3×10^{-10}	2.4×10^{-10}
$^3P_0 - {}^3P_1$	1.8×10^{-5}	2.1×10^{-12}

Table 7.2: *Left: level designations, energies, statistical weights of hyperfine structure of O I $2p^4\ ^3P$. Right: level designations of the forbidden hyperfine transition, Einstein A coefficient, collision rate coefficient for collisions with molecular hydrogen at 100 K.*

b) List the processes that impact the level population numbers of these fine structure states.

The critical density for a given downward transitions $u \to l$ is given by

$$n_{\rm cr} = \frac{A_{ul}}{q_{ul}}. \tag{7.34}$$

See section 15.3 for the physical meaning of this quantity.

c) Calculate the critical density of the [O I] 63 μm line transition and compare this value to a typical molecular cloud. Do you expect the level populations to be in Local Thermodynamic Equilibrium in that environment?

d) Predict the LTE total intensity in erg $cm^{-2}\ s^{-1}\ sr^{-1}$ of the [O I] 63 μm line from an optically thin slab of gas with an oxygen column density $N(\mathrm{O}) = 10^{12}$ cm^{-2} and constant temperature of 100 K. (Tip: consult Eq. 7.16).

Exercise 7.5

We consider an atomic model with only two bound levels, in which the lower and upper level are respectively labeled 1 and 2.

a) Give the probability ϵ that a photon that causes a radiative excitation from level 1 is destroyed by a collisional de-excitation event. This probability is generally referred to as the *photon destruction probability*.

b) Show that in TE

$$\epsilon = \frac{C_{21}}{C_{21} + A_{21}/\left(1 - e^{-h\nu_{12}/kT}\right)} \tag{7.35}$$

Exercise 7.6

a) Show that for the two level atom, line source function Eq. (7.28) can be written as

$$S_\nu^\ell = \frac{\epsilon' B_\nu + \bar{J}_{12}}{1+\epsilon'}, \tag{7.36}$$

where

$$\epsilon' = \frac{C_{21}}{A_{21}} \left(1 - e^{-h\nu_{12}/kT}\right). \tag{7.37}$$

b) Show that

$$S_\nu^\ell = \epsilon\, B_\nu + (1-\epsilon)\, \bar{J}_{12}, \tag{7.38}$$

where

$$\epsilon = \frac{\epsilon'}{1+\epsilon'}. \tag{7.39}$$

One can readily see that expressions Eq. (7.39) and (7.35) are identical. Source function (7.36) contains a thermal emissivity term $\epsilon'\, B_\nu$. In essence, it represents photons created by collisional excitation followed by radiative de-excitations. The term ϵ' in the denominator is a sink term that represents photons that are destroyed by a collisional de-excitation following a photo-excitation. $\bar{J}_{12}$ is the non-coherent scattering emissivity term. It is non-coherent (i.e. the emitted photon has forgotten about the nature of the absorbed photon) as we have explicitly adopted complete redistribution of the line photons at the very start of our derivations. See section 14.2 for a more extended discussion of the two-level problem.

8

Continuum processes

Opacities

In this chapter we treat processes that give rise to *continuum* extinction and emission. An overview of the behavior of the mass extinction coefficient χ'_ν (see Eq. 4.5) as function of temperature and density is given in Figure 8.1. The unit of χ'_ν is $\mathrm{cm^2 gr^{-1}}$, i.e. it provides the total cross section per gram of material. In the figure the density is chosen such that it represents common environments characterized by a temperature T. For instance, in the densest fragments of an interstellar molecular cloud the particle density may reach $n \sim 10^7\,\mathrm{cm^{-1}}$, hence $\rho \sim 10^{-19}\,\mathrm{gr\,cm^{-3}}$. In the atmosphere of a giant star of 5 000 K a typical density is $10^{-8}\,\mathrm{gr\,cm^{-3}}$.

The mass extinction coefficient is dominated by dust grains at temperatures up to about 1500-2000 K and by molecules and atoms above that temperature range. Between temperatures of about 400 and 1500 K the grains are composed of mainly silicates, iron, and their compounds. Between 400 and 175 K, volatile organics such as methane (CH_4), ethane (C_2H_6) and ethanol (C_2H_5OH) are also present. Below 175 K, ice grains are present along with the other species and are actually dominating. In this low temperature range the Rosseland mean opacity κ_R (Eq. 4.66) increases as T^2. Between 175 K and about 1700 K the opacity is dominated by silicates and iron grains. The sudden dip at about 1700 K is caused by the evaporation of the silicate and iron grains; above that temperature and up to about 3500 K molecular opacity dominates with important contributors being H_2O, CH_4, NH_3, and CO. At a few thousand Kelvin and up, the hydride ion H^- gains in importance, dominating the opacity above 3500 K up to about 8000 K. The Rosseland opacity in this regime increases as $\rho^{1/2}\,T^{9/2}$. At 8000 K bound-free and free-free opacity of hydrogen takes over with Thomson scattering on free electrons gaining importance. Above 10000 K scattering on free electrons becomes the main opacity source.

Calculating the Rosseland mean opacity for a gas mixture is very challenging as it requires knowledge of the individual opacities of tens to hundreds of millions of atomic and, particularly, molecular transitions from neutral and ionized stages. Much of the pioniering work in this field has been done in relation to nuclear bomb experiments after the second world war.

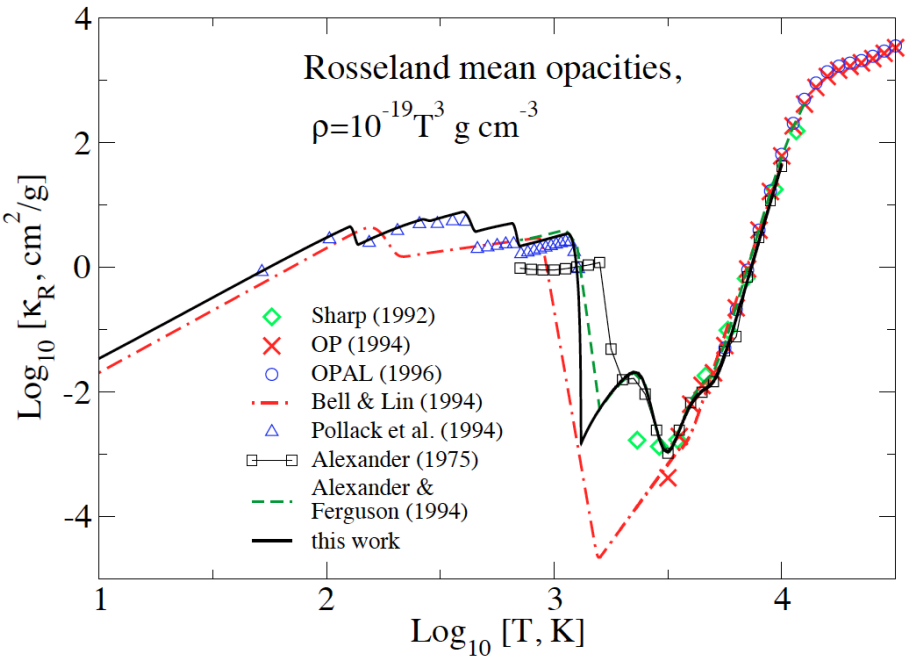


Figure 8.1: *The Rosseland mean opacities in* $cm^2\ gr^{-1}$ *calculated for temperatures between 1 K and* 10^5 *K and gas densities* $\rho = 10^{-19}\ T^3\ gr\,cm^{-3}$*. The adopted C/O abundance is 0.43. Dust grains dominate the opacity up to a temperature of about 1500-2000 K; atoms dominate beyond 4000 K. The dust grains consist of a mix of silicates, organics, amorphous ice, FeS, and iron. The ice grains evaporate at about 160 K. The results of Semenov et al. are compared to opacity tables provided by the Opacity Project (OP; 1994, crosses), the Opacity Project at Livermore (OPAL; 1996, circles) and other, smaller, projects. From: Semenov et al. 2003.*

Later on, astrophysics benefitted from this data. Notable large efforts have been the Opacity Project (OP) and the Opacity Project at Livermore (OPAL). Results of these projects are compared to the work of Semenov et al. (2003) in the overlapping part of the parameter domain in Fig. 8.1.

In the remainder of this chapter we focus on atomic and ionic material, i.e., we refrain from a discussion of molecular and dust opacities.

8.1 Bound-free transitions

Similar to there being five bound-bound processes, we identify also five bound-free processes:

1. *photoionization*
   requires: photons of sufficient energy

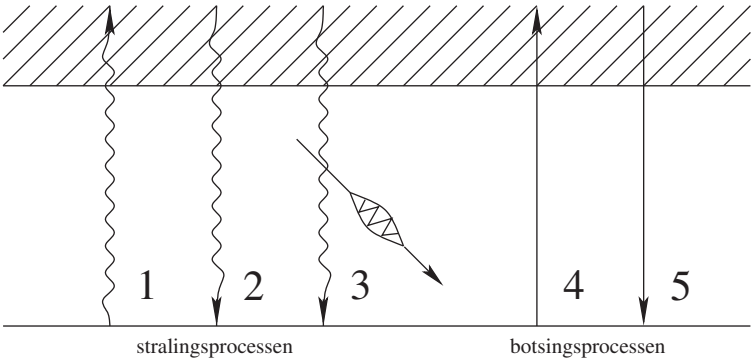


Figure 8.2: *Schematic overview of the five possible bound-free processes.*

2. *spontaneous photo-recombination*
   requires: free electron that can be captured

3. *stimulated photo-recombination*
   requires: free electron that can be captured and a photon to induce the process

4. *collisional ionization*
   requires: passing particle with sufficient energy

5. *collisional recombination*
   requires: passing particle and a free electron that can be captured

The processes 1 and 4 require an energy (from photon or collision) $E > E_I - E_i = E_\circ$, where E_i is the energy of the bound level i and E_I is the ionization energy. The relevant extinction coefficient has a threshold value, $\nu_\circ = E_\circ/h$ for photoionization, and $v_\circ = (2E_\circ/m_\mathrm{e})^{1/2}$ for collisional ionization, below which the process can not take place.

The latter process is a three particle collision and is therefore, as a rule, rare.

Photoionization

In a *photoionization* a particle in a bound energy state i is ionized, through absorption of a photon. We refer to the ionized state as κ.

Say, the extinction coefficient for photoionization or *photoionization cross-section* per particle in energy level i is given by $\alpha_{i\kappa}(\nu)$, which has dimension cm^2. If the energy of a photon is less than what is required to remove the electron from the atom or ion, i.e. $\nu < \nu_\circ$, then $\alpha_{i\kappa}(\nu) = 0$. The rate of photoionization follows from the product of this cross-section and the number of photons that are available at each frequency $\nu \geq \nu_\circ$, i.e. $4\pi J_\nu/h\nu$.

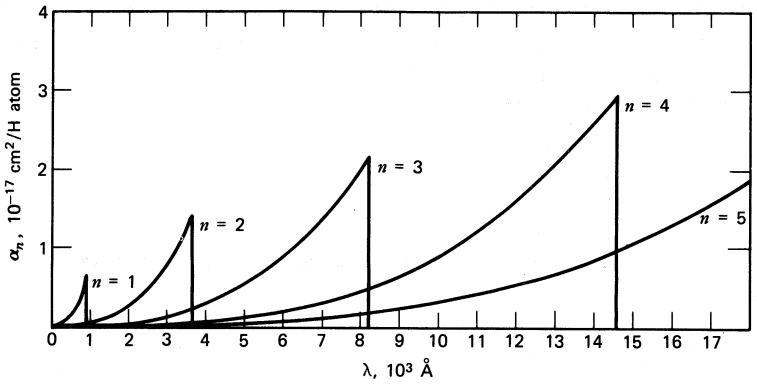


Figure 8.3: *The bound-free extinction coefficient per particle for hydrogen atoms, for different principal quantum numbers n. Although $\alpha_\nu^{\rm bf}(n)$ increases with n one should be reminded that the particle number density of excited levels, n_n, usually is a strongly decreasing function of n (see for instance eq. 6.21) such that the behaviour of the linear extinction coefficient as a function of n can be more complex (see for instance fig 8.6). From: Gray (1992).*

From this consideration it follows that the number of photoionizations per second per cm^3 is given by

$$n_i \, R_{i\kappa} = n_i \, 4\pi \int_{\nu_\circ}^{\infty} \alpha_{i\kappa}(\nu) \, \frac{J_\nu}{h\nu} \, d\nu \tag{8.1}$$

For hydrogenic ions (H I, He II, Li III, *etcetera*) the extinction coefficient per particle for photoionization (without correction for induced recombinations) from a level of principal quantum number n, for $\nu \geq \nu_\circ(n, Z)$, is given by

$$\alpha_\nu^{\rm bf}(n, Z) = \left(\frac{64\pi^4 m_{\rm e} e^{10}}{3\sqrt{3}\, ch^6} \right) \frac{Z^4}{n^5} \, g_{\rm II}(\nu, n) \, \nu^{-3} \tag{8.2}$$

where Z is the charge of the ion ($Z = 1$ for hydrogen) and $g_{\rm II}$ is the bound-free Gaunt factor. Close to the ionization edge $g_{\rm II} \sim 1$. For this type of ion the excitation energy is given by $E(n, Z) = \mathcal{R}Z^2(1 - 1/n^2)$, where $\mathcal{R}$ is the Rydberg constant for hydrogen, such that the edge frequency $\nu_\circ(n, Z) = (\mathcal{R}Z^2/hn^2)$. Using this relation and substituting constants yields

$$\alpha_\nu^{\rm bf}(n, Z) = 7.91 \times 10^{-18} \frac{n}{Z^2} \, g_{\rm II}(\nu, n) \left(\frac{\nu_\circ}{\nu} \right)^3 \tag{8.3}$$

The extinction is thus proportional to ν^{-3} for $\nu > \nu_\circ$. For more complicated ions, that have more valance electrons (for instance Fe I which has an electron shell that is half filled and that provides a multitude of valance electrons and valance cavities), the smooth decline is disrupted by resonances which produce all sorts of peaks in $\alpha_{i\kappa}(\nu)$ that need to be determined in experiments.

Spontaneous recombination

In a *spontaneous recombination* an electron is captured by a particle in an ionized state κ and a photon is emitted that carries with it the recombination energy. We may derive the number of spontaneous recombinations per second per cm^3 using a detailed balance argument.

In TE the number of photoionizations, at each frequency, must be in balance with the number of radiative recombinations. As in this most strict of all equilibria $J_\nu = B_\nu$, it follows that in TE the total number of radiative recombinations per cm^3 per second is

$$n_\kappa^* \, R_{\kappa i} = n_i^* \, R_{i\kappa} = n_i^* \, 4\pi \int_{\nu_\circ}^{\infty} \alpha_{i\kappa}(\nu) \, \frac{B_\nu}{h\nu} \, d\nu \tag{8.4}$$

The superscript $*$ indicates that the level occupations are according to the Boltzmann and Saha equilibrium (see for instance eq. 6.26). The total number of radiative recombinations is the sum of spontaneous and induced recombinations, and we may split these two contributions using the correction term for stimulated emission

$$\begin{aligned} n_\kappa^* \, R_{\kappa i} &= n_\kappa \, R_{\kappa i}^{\rm spon} + n_\kappa \, R_{\kappa i}^{\rm stim} \qquad (8.5) \\ &= n_i^* \, 4\pi \left[\int_{\nu_\circ}^{\infty} \frac{\alpha_{i\kappa}(\nu)}{h\nu} \, B_\nu \left(1 - e^{-h\nu/kT}\right) \, d\nu + \int_{\nu_\circ}^{\infty} \frac{\alpha_{i\kappa}(\nu)}{h\nu} \, B_\nu \, e^{-h\nu/kT} \, d\nu \right] \end{aligned}$$

Using the definition of the Planck function (eq. 6.4) we find for the spontaneous recombinations

$$n_\kappa^* \, R_{\kappa i}^{\rm spon} = n_i^* \, 4\pi \int_{\nu_\circ}^{\infty} \frac{\alpha_{i\kappa}(\nu)}{h\nu} \, \frac{2h\nu^3}{c^2} \, e^{-h\nu/kT} \, d\nu \tag{8.6}$$

For a spontaneous recombination to occur what is required is a passing free electron that can be captured. What is *not* required is knowledge of possible deviations from TE of the local radiation field. Even if TE (or LTE) is not valid, the rate $R_{\kappa i}^{\rm spon}$ should still hold per ion, as long as the velocity distribution of the free electrons is according to Maxwell. We therefore only need to rescale eq. (8.6) from the TE ion density to the real ion density n_κ to get a general expression for the number of spontaneous radiative recombinations per cm^3 per second:

$$n_\kappa \, R_{\kappa i}^{\rm spon} = n_\kappa \left(\frac{n_i}{n_\kappa} \right)^* 4\pi \int_{\nu_\circ}^{\infty} \frac{\alpha_{i\kappa}(\nu)}{h\nu} \, \frac{2h\nu^3}{c^2} \, e^{-h\nu/kT} \, d\nu \tag{8.7}$$

This expression states that the emission coefficient for spontaneous recombination, per ion, "feels" the Planck function (which represents the Maxwell velocity distribution) and not the real radiation field. The latter may obviously have caused the ionization into state κ, however, this dependence is locked up in n_κ (and *not* in $R_{\kappa i}^{\rm spon}$). It may be clear that this recombination process is less "spontaneous" as is a spontaneous de-excitation between two bound levels; after all, the ion needs to wait for a free electron that can be captured. The dependence of $R_{\kappa i}^{\rm spon}$ on this free electron is hidden in the TE ratio $(n_i/n_\kappa)^*$ and is explicitly given by the Saha-Boltzmann equation (6.26). Substitution yields

$$\begin{aligned} n_\kappa \, R_{\kappa i}^{\rm spon} &= n_\kappa n_{\rm e} \, \Phi_{i\kappa}(T) \, 4\pi \int_{\nu_\circ}^{\infty} \frac{\alpha_{i\kappa}(\nu)}{h\nu} \, \frac{2h\nu^3}{c^2} \, e^{-h\nu/kT} \, d\nu \\ &\equiv n_\kappa \, n_{\rm e} \, \alpha_{i,\kappa-1}^{\rm RR}(T) \qquad (8.8) \end{aligned}$$

where $\alpha^{\rm RR}_{i,\kappa-1}(T)$ is the *recombination coefficient* in cm^3 per second into level i of ion $\kappa-1$. This function decreases with increasing temperature, as it is more difficult for an ion to capture a faster moving electron (see e.g. table 18.1). Equation (8.8) shows that the number of spontaneous recombinations scales with the product $n_\kappa n_{\rm e}$, as expected. The recombination coefficient does not account for stimulated recombinations. These are however in most cases unimportant compared to the spontaneous recombinations (see for instance eq. 8.5 for the case $J_\nu = B_\nu$ and realize that $\nu_\circ$ usually is quite large).

Stimulated recombination

The same argument that applies to the process of spontaneous recombination also holds for *stimulated recombinations* or *induced recombinations*. In TE the number of stimulated recombinations per cm^3 per second is (see eq. 8.5)

$$n^*_\kappa\, R^{\rm stim}_{\kappa i} = n^*_i\, 4\pi \int_{\nu_\circ}^{\infty} \frac{\alpha_{i\kappa}(\nu)}{h\nu}\, B_\nu\, e^{-h\nu/kT}\, d\nu \tag{8.9}$$

Rescaling of the TE ion density n^*_κ to the real ion density n_κ, and substituting the TE radiation field B_ν for the real radiation field J_ν, yields

$$n_\kappa\, R^{\rm stim}_{\kappa i} = n_\kappa \left(\frac{n_i}{n_\kappa}\right)^* 4\pi \int_{\nu_\circ}^{\infty} \frac{\alpha_{i\kappa}(\nu)}{h\nu}\, J_\nu\, e^{-h\nu/kT}\, d\nu \tag{8.10}$$

The end result for the *total number of radiative recombinations*, per cm^3 per second, therefore is

$$n_\kappa\, R_{\kappa,i} = n_\kappa \left(\frac{n_i}{n_\kappa}\right)^* 4\pi \int_{\nu_\circ}^{\infty} \frac{\alpha_{i\kappa}(\nu)}{h\nu} \left[\frac{2h\nu^3}{c^2} + J_\nu\right] e^{-h\nu/kT}\, d\nu \tag{8.11}$$

Collisional ionization

The number of *collisional ionizations* per cm^3 per second follows from eq. (7.7), where the threshold velocity $v_\circ$ now corresponds to the minimum energy $E_\circ = \frac{1}{2}mv^2_\circ = h\nu_\circ$ of the electron needed for collisional ionization. We get

$$n_i\, C_{i\kappa} = n_i\, n_{\rm e} \int_{v_\circ}^{\infty} \sigma_{i\kappa}(v)\, f(v)\, v\, dv \equiv n_i\, n_{\rm e}\, q_{i\kappa}(T) \tag{8.12}$$

For general remarks concerning the collisional process we refer to § 7.1. Here we only give the approximate formula for $q_{i\kappa}(T)$ based on a dipole approximation, derived by Van Regemorter,

$$q_{i\kappa}(T) \approx 2.7\,\zeta \left(\frac{E_\circ}{kT}\right)^{-2} T^{-3/2}\, e^{-E_\circ/kT} \tag{8.13}$$

where ζ is the number of electrons in the outer shell.

Collisional recombination

The rate of collisional recombinations can also be obtained using a detailed balance argument. In KE

$$n_i^* \, C_{i\kappa} = n_\kappa^* \, C_{\kappa i} \tag{8.14}$$

such that the number of collisional recombinations per cm^3 per second is

$$n_\kappa \, C_{\kappa i} = n_\kappa \left(\frac{n_i}{n_\kappa} \right)^* C_{i\kappa} = n_\kappa \, n_{\rm e}^2 \, \Phi_{i\kappa}(T) \, q_{i\kappa}(T) \tag{8.15}$$

The equation shows clearly that this is a three particle process, i.e. the process scales as $n_i n_{\rm e}^2$. The rate of collisional recombinations is much less sensitive to temperature as is that of collisional ionizations, because, similar to collisional processes between bound energy levels (see § 7.1), the exponential term drops out. Collisional recombinations are more likely to occur for energy levels high up in the atom/ion because of the low threshold energy $E_\circ$.

Collisional coupling of high levels with the continuum

The collisional coupling between high lying levels and the continuum is even stronger than that between two bound energy levels close to the ionization edge (see § 7.1). This is so because $q_{i\kappa}(T)$ for given temperature increases more rapidly with decreasing E_0 as does q_{lu} (see eq. 7.8). If the electron density decreases, for instance high up in the atmosphere or in a stellar wind, spontaneous recombination and spontaneous de-excitation processes will win over collisional processes for an increasing number of levels. Consequently, these levels will cascade or "rain down" into the ground level. However, for levels that are extremely high up in the energy level diagram (i.e. those close to the ionization threshold) the collisional processes will continue to dominate and therefore the populations will remain in LTE (relative to the continuum) and will be given by eq. (6.26).

8.2 Free-free transitions

Radiation that is the result of the negative (see below) acceleration of a charge in the Coulomb field of another charge is known as *free-free radiation*. In a dipole approximation free-free radiation only occurs if the two particles are not identical. So, it does not happen for electron-electron or ion-ion interactions, as in those cases the electric dipole moment $\mathbf{d} \equiv \sum q_i \mathbf{r}_i$ is proportional to the center of mass $\sum m_i \mathbf{r}_i$, which is a conserved quantity. This causes the emitted power, given by Larmor's law

$$P = \frac{2\ddot{d}^2}{3c^3} \tag{8.16}$$

to be zero. In electron-ion interactions the electrons are the prime radiators as the relative acceleration $\ddot{\mathbf{r}}$ is inversely proportional to the mass of the particles squared, while the charges are typically about equal. Free-free radiation is also referred to as *bremsstrahlung*. Bremsstrahlung is the German for 'braking radiation', since the emitting electron-ion pair are being braked rather than accelerated.

We give, without derivation, the result for the free-free extinction coefficient $\chi_\nu^{\rm ff}$ per cm for particles of element k in ionization state j

$$\chi_\nu^{\rm ff} = \frac{4e^6}{3hc}\left(\frac{2\pi}{3km_{\rm e}^3}\right)^{1/2} \overline{g}_{\rm III}(\nu,T)\,\frac{Z_{jk}^2}{T^{1/2}\nu^3}\,n_{\rm e}\,N_{jk}\left(1-e^{-h\nu/kT}\right) \tag{8.17}$$

where $\overline{g}_{\rm III}$ is the velocity mean for the Gaunt factor of free-free processes, which gives the quantum mechanical correction to the classic result; $\overline{g}_{\rm III}$ is dimensionless and of order unity (but see eq. 17.39 for its behavior in the radio regime). N_{jk} is the density of ions of charge Z_{jk}. For H II and He II, $Z_{jk}=1$; for He III, $Z_{jk}=2$. Evaluation of all constants yields a cgs value 3.692×10^8.

In TE the free-free source function is given by the Planck function B_ν(T). This is also true if TE is not valid, provided that the velocity distribution of the particles is given by the Maxwell equation, because in free-free processes kinetic and radiative energy is exchanged fully. In each free-free emission a photon is created out of the "thermal pool of particles" that has no prior knowledge about anything; likewise, in a free-free absorption all information about the absorbed photon will be lost. In other words: a free-free process is a *pure* collisional process.

In the Wien approximation $(h\nu/kT \gg 1)$ the correction term for stimulated emissions may be neglected. One is left with

$$\chi_\nu^{\rm ff} = 3.692\times10^8\,\overline{g}_{\rm III}(\nu,T)\,\frac{Z_{jk}^2}{T^{1/2}\nu^3}\,n_{\rm e}\,N_{jk} \tag{8.18}$$

i.e. $\chi_\nu^{\rm ff}\propto\nu^{-3}$. In the Rayleigh-Jeans regime $(h\nu/kT\ll1)$

$$\chi_\nu^{\rm ff} = 1.772\times10^{-2}\,\overline{g}_{\rm III}(\nu,T)\,\frac{Z_{jk}^2}{T^{3/2}\nu^2}\,n_{\rm e}\,N_{jk} \tag{8.19}$$

which shows a frequency dependence $\chi_\nu^{\rm ff}\propto\nu^{-2}\propto\lambda^2$.

8.3 H^- absorption

Hydrogen, because of its large polarizability, can form a negative ion consisting of a proton and two electrons. The H^- or hydride ion has a single bound state with a binding energy of 0.754 eV. The wavelength corresponding to this energy is 16444 Å. Therefore, it has no lines, only bound-free and free-free transitions. For bound-free processes

$$H^- + h\nu \rightleftharpoons H + e^-(v) \tag{8.20}$$

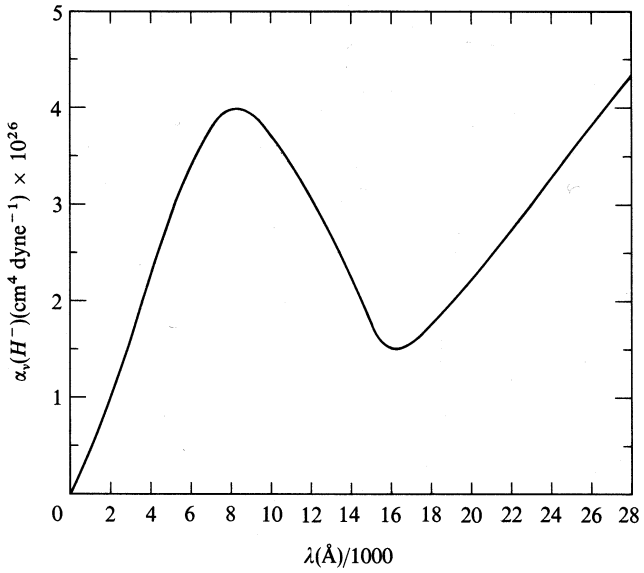


Figure 8.4: *Extinction coefficient (in 10^{26} cm^2) of the* H$^-$ *ion, per neutral H atom and per unit of electron pressure $p_\mathrm{e} = n_\mathrm{e}kT$, in LTE at 6300 K. The bound-free extinction has a maximum at $\sim$ 8000 Å. The free-free extinction is proportional to λ^2. The sum shows a maximum at 8500 Å and a minimum at 1.6 μm. From Mihalas (1978).*

where $1/2\, m_\mathrm{e}v^2 = h\nu - 0.754$ eV. For free-free processes

$$\mathrm{H}^- + e^-(v) + h\nu \rightleftharpoons \mathrm{H}^- + e^-(v') \tag{8.21}$$

where $1/2\, m_\mathrm{e}v'^2 = h\nu + 1/2\, m_\mathrm{e}v^2$. The bound-free continuum does not have a sharp ionization edge at 16444 Å, but a broad peak at lower wavelength reaching a maximum at $\sim$ 8500 Å (see figure 8.4).

Because of its low binding energy H$^-$ can only exist in relatively cool stars (up to stars slightly hotter than the sun). In the photosphere of the sun and in that of cooler stars the H$^-$ processes are the dominant source of visual and infrared extinction. The extra electrons are supplied by elements such as Na, Mg, Al, Si, Ca, en Fe, that are relatively abundant and have a much lower ionization energy than hydrogen (see also § 6.5). In extremely cool stars, brown dwarfs and hot gaseoous planets H$^-$ will not be the dominant source of extinction, because of a lack of free electrons, but it may be one of the contributing sources.

The identification of H$^-$ as the most important source of extinction in cool stars by Pannekoek in 1930 and 1935 and Wildt in 1939 and 1942 was viewed as a fundamental breakthrough. To illustrate this: Eddington, in his book 'The internal constitution of the stars' in 1926 saw the unknown extinction and the (at that time) also unknown internal source of energy of the stars as the *only* remaining problems in the physics of stars!

In TE (or LTE) the number of H$^-$ ions is given by the Saha equation (6.24). Usually one

formulates this using an explicit reference to the electron pressure $p_\mathrm{e} = n_\mathrm{e} kT$, such that $N_{\mathrm{H}^-} = N_\mathrm{H}\, p_\mathrm{e} \breve{\Phi}(T)$. This yields for the H^- extinction

$$\chi^*_\nu(\mathrm{H}^-) = \alpha^\mathrm{bf}_\nu(\mathrm{H}^-)\, N_\mathrm{H}\, p_\mathrm{e}\, \breve{\Phi}(T) \left[1 - e^{-h\nu/kT}\right] \tag{8.22}$$

where $\alpha^\mathrm{bf}_\nu(\mathrm{H}^-)$ is the extinction coefficient per cm^2. For the free-free extinction $\alpha^\mathrm{ff}_\nu(\mathrm{H}^-)$ a similar expression can be given. Fitting formulas for $\alpha^\mathrm{bf}_\nu(\mathrm{H}^-)$ and $\alpha^\mathrm{ff}_\nu(\mathrm{H}^-)$ are discussed in, for instance, Gray.

8.4 Elastic scattering by electrons

We discuss the scattering of radiation by free and bound electrons in non-relativistic media, such that the scattering is elastic, implying the frequency and energy remain unchanged and only changes in direction occur.

We do not discuss extinction or emission due to electrons being accelerated in a magnetic field, i.e. cyclotron radiation (in case the electrons have non-relativistic velocities) and synchrotron radiation (in case the electrons do have relativistic velocities).

Thomson scattering

In the limit of low photon energy, i.e. $h\nu \ll m_\mathrm{e} c^2$, the scattering by *free electrons* of low energy will be *elastic* or *coherent*. This process is referred to as *Thomson scattering*. For high photon energies Compton shifts will occur because of the change of momentum and energy between photon and electron. If the scattering by free electrons involves high energy electrons one needs to apply relativistic corrections; this process is referred to as inverse Compton scattering. We will only concentrate on Thomson scattering.

The extinction coefficient per free electron is given by

$$\sigma_\mathrm{T} = \frac{8\pi e^4}{3 m_\mathrm{e}^2\, c^4} = 6.652 \times 10^{-25}\ \ \mathrm{cm}^2 \tag{8.23}$$

and is independent of the frequency of the incident radiation. The result can be obtained using a classical description in which the electron is viewed as a free particle that vibrates (without being damped) in a passing electromagnetic wave.

The linear extinction coefficient is given by

$$\chi^\mathrm{e} = \sigma_\mathrm{T}\, n_\mathrm{e} \tag{8.24}$$

where n_e is the electron density. Note that no subscript ν is given as the process is independent of frequency. Thomson scattering is the dominant source of continuous extinction in the

atmospheres of early-type stars where hydrogen is almost fully ionized. The much larger value of the bound-free extinction, that is of the order of 10^{-17} cm^2 per particle (see eq. 8.3) usually wins from scattering by free electrons if hydrogen is only partly ionized.

Specifying the mass fraction of hydrogen in a gas by X, and the ionization fraction of hydrogen by q, we find for the contribution of hydrogen to the mass extinction coefficient by Thomson scattering

$$\chi'_{\rm e}({\rm H}) = \frac{\sigma_{\rm T}}{m_{\rm H}}\, X\, q = 0.40\, X\, q \qquad {\rm cm}^2/{\rm gr} \tag{8.25}$$

where $m_{\rm H}$ is the mass of the hydrogen atom. For a gas of solar abundances $X = 0.747$. If in this gas hydrogen in fully ionized (q = 1) we find $\chi'_{\rm e}({\rm H}) = 0.30$ cm^2/gr.

Rayleigh scattering

Photons can also be scattered by *bound electrons* in for instance an atom or molecule or for any other particle that is much smaller than the wavelength of the incident light. In the classic picture this bound electron will have a resonance frequency that is equal to the eigen frequency ν_{lu} of a harmonically bound electron. In the limit in which the frequency of the incident radiation $\nu \ll \nu_{lu}$, we have Rayleigh scattering. In this case the cross section per valance electron per resonance line is given by

$$\sigma_{\rm Ray}(\nu) = \sigma_{\rm T}\, f_{lu}\, \frac{\nu^4}{(\nu^2 - \nu_{lu}^2)^2} \simeq \sigma_{\rm T}\, f_{lu} \left(\frac{\nu}{\nu_{lu}}\right)^4 \tag{8.26}$$

where the oscillator strength gives the quantum mechanical correction of the resonance strength compared to the classic value.

The Rayleigh scattering cross section of a gas is also related to its index of refraction m:

$$\sigma_{\rm Ray} = \frac{8\pi^3}{3c^4} \left(\frac{m^2 - 1}{n}\right)^2 \nu^4 \tag{8.27}$$

where n is the number density of scattering centers per cm^{-3}.

The ratio between the cross section for Rayleigh and Thomson scattering is $\sigma_{\rm Ray}(\nu)/\sigma_{\rm T} \simeq f_{lu}(\nu/\nu_{lu})^4 \ll 1$. Rayleigh scattering can only be a significant contribution to the total extinction coefficient if the medium contains only very few free electrons, and if other sources of extinction do not spoil things. This is the case in G- and K-type stars, and specifically those of population III. The latter have very low abundances of metals, which are the most important donors of free electrons in G and K stars (see § 6.5). This implies that H$^-$ as a source of extinction is relatively unimportant. In G- and K-type stars hydrogen is almost completely in the ground state. The resonance frequencies of the Lyman series lines (i.e. those originating from the ground level) are in the ultraviolet part of the spectrum, such that photons in the visual part of the spectrum can interact with these transitions through the Rayleigh scattering

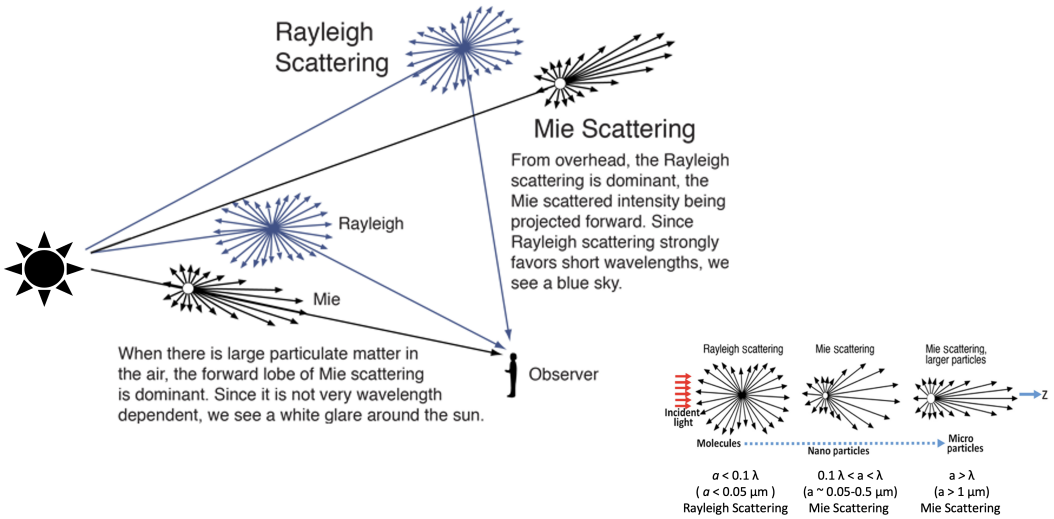


Figure 8.5: *Blue light from the sun (to the left of the figure) is much more efficiently scattered by molecules in the Earth atmosphere than red light. Blue photons from beams of solar light that would not directly reach our eyes are thus more often scattered in the direction of our eyes. This is why the sky appears blue. Also shown in the figure are scatterings of solar photons on relatively large solid state particles in the Earth atmosphere. Such scatterings are not very wavelength dependent but occur mostly in almost forward directions. When such particles are dispersed in the air one sees a white glare around the sun.*

mechanism. In the very low temperature M stars, brown dwarfs and (exo)planets, molecular Rayleigh scattering will dominate because hydrogen, the most abundant element, is mainly present in molecular form. For H_2 a convenient formula for $\sigma_{\rm Ray}(H_2)$ is given by Dalgarno & Williams (1962), see also Seager (2010). The cross section of Rayleigh scattering per H_2 molecule is comparable to that of atomic hydrogen.

Rayleigh scattering of sunlight by N_2 and O_2 is important in the earth atmosphere and causes the sky to have a blue color and the setting sun to become redder. A useful numerical formula for the Rayleigh scattering cross section for Earth's atmosphere is given by Bates (1994). In the visible spectral range, from 4000 Å (violet) tot 8000 Å (red), the extinction coefficient of Rayleigh scattering decreases by a factor $2^4 = 16$. Therefore, blue photons will experience significantly more scattering than do red photons. The probability that blue light is scattered out of the sun-eye direction is thus much larger than for red light. At sunset, when the column mass (see eq. 9.34) in the line-of-sight becomes large, this effect causes the sun to become redder. Scattered blue photons that originally traveled in a direction that would not reach our eyes have a larger probability of being scattered into a line-of-sight that reaches our eyes. The effect is that the observer sees blue photons from arbitrary directions. This explains the blue color of the sky (see Fig. 8.5).

8.5 Comparison of the main sources of extinction

Figure 8.6 shows the relative importance of continuous extinction by H I, H^-, and He^- in the photosphere of three types of dwarf stars, as a function of wavelength. The contribution of Thomson scattering is not given as for these three cases it is negligibly small (the stars are too cool to have large numbers of free electrons). We first make an estimate to show that this is so.

The extinction coefficient given on the vertical axis (in units of 10^{26} cm^2) is measured per neutral hydrogen atom for an arbitrary excitation state (distributed according to Boltzmann), and normalized to the electron pressure. The latter is done because free-free and H^- extinction scale with p_e (see §§ 8.2 and 8.3 respectively). The linear extinction coefficient follows from the values given in figure 8.6, which we will call $\kappa_0(\nu)$, via $\chi_\nu = 10^{-26}\kappa_\circ(\nu)\, N_0\, p_e$ cm^{-1}. In the unit of κ_0 the extinction by Thomson scattering is $10^{26}\sigma_T\, n_e/N_0\, p_e = (66.5/p_e)\, n_e/N_0$. The value of the electron pressure, as given in the three panels, is of the order 10^2-10^3 dyn cm^{-2}. If we assume a pure hydrogen gas, the ratio $n_e/N_0 = N_+/N_0$ is derived immediately from the Saha ionization equation 6.24. As hydrogen is almost completely neutral in the three examples, this ratio $\ll 1$. This shows that the extinction coefficient of Thomson scattering in the unit applied in the figure $\ll 1$, i.e. it is negligible compared to the other sources of extinction.

The top panel in figure 8.6 shows the extinction in a model of a solar-type dwarf (an early-G star). The absorption is dominated by H^- (compare with figure 8.4). The bound-free absorption by hydrogen is significantly larger in the Balmer continuum compared to the Paschen continuum, whereas the cross-section for ionization $\alpha_\nu^{\rm bf}$ (eq. 8.3 and figure 8.3) shows an increase with excitation level. The reason is in the temperature dependence of the level population numbers of the exited states: there are much less H-atoms in level n=3 (Paschen) compared to atoms in level n=2 (Balmer).

The mid panel is for a late-A dwarf. The bound-free and free-free contributions of neutral hydrogen are the most important. Of these two the relative importance of free-free extinction increases for increasing wavelength as in the optical and near-infrared $\chi_\nu^{\rm ff} \propto \lambda^3$ (see eq. 8.17). This explains the increase of the H-absorption with wavelength. The negative hydrogen ion still adds somewhat to the total extinction.

The bottom panel shows the extinction for a late-B dwarf. The H^--ion is no longer contributing in any significant way. The absorption is dominated by hydrogen. In the optical the bound-free transitions dominate, while in the near-infrared free-free processes are the most important.

Other sources of extinction

Figure 8.6 does not show the extinction behaviour at ultraviolet wavelengths. In this part of the spectrum the situation is more complex as also bound-free extinction of heavy elements (such as Mg, Al, Si, and C) start to play a role. For the hot O- and early B-type stars the bound-free edges of helium are an important source of absorption in the ultraviolet. The He I continuum starts at 504 Å; that of He II at 228 Å.

Stars cooler than the sun feature several molecular ions, such as H_2^-, CN^-, C_2^- and H_2O^- which contribute strongly to the continuous extinction. Also Rayleigh scattering by hydrogen atoms and H_2 molecules plays a role. For the extremely cool M-type stars molecular bands, notably those of TiO, are strong sources of continuous absorption (see chapter 2).

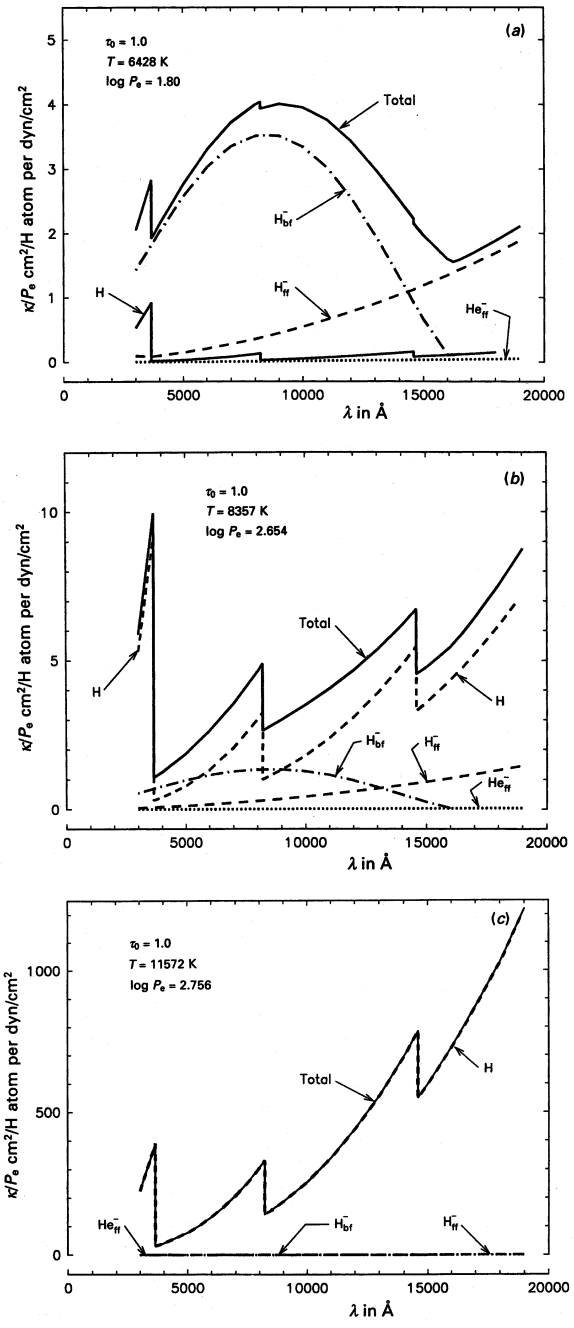


Figure 8.6: *The extinction coefficient (in 10^{-26} cm^2; note that the scaling by a factor 10^{+26} is not indicated on the vertical axis) for hydrogen and helium for different continuum processes, per neutral hydrogen atom and per unit electron pressure measured at an optical depth $\tau_\circ = 1$ for continuum radiation at a wavelength of 5000 Å. Panel (a) is for a sun-like dwarf; (b) for a late-A dwarf, and (c) for a late-B dwarf. From Gray (1992).*

Exercise 8.1

Give an expression for the total linear extinction coefficient for bound-free processes, ${\chi_\nu}^{\rm bf}$, in a pure hydrogen gas in LTE.

Exercise 8.2

Show that the free-free emission coefficient is given by

$$\begin{aligned} \eta_\nu^{\rm ff} &= \frac{8e^6}{3c^3}\left(\frac{2\pi}{3k\,m_{\rm e}^3}\right)^{1/2} \overline{g}_{\rm III}(\nu,T)\,\frac{Z_{jk}^2}{T^{1/2}}\,n_{\rm e}\,N_{jk}\,e^{-h\nu/kT} \\ &= 5.443\times10^{-39}\,T^{-1/2}\,n_{\rm e}\,n_{\rm p}\,e^{-h\nu/kT}, \end{aligned} \tag{8.28}$$

where the latter equation assumes an ionized hydrogen gas and $\overline{g}_{\rm III}(\nu,T)=1$

Exercise 8.3

This could be a nice exam question. From the dispersion of radial velocities of galaxies in the Coma cluster and applying the virial theorem, one finds a total mass for the Coma cluster of $3.3\times10^{15}M_\odot$. The total mass of all stars in all galaxies of the Coma cluster can be derived from the total UV to infrared luminosity of the entire system and is $1.5\times10^{13}M_\odot$. The Coma cluster also contains hot gas, with $T_{\rm gas}=8.8\times10^7$ K. In this exercise we investigate whether this hot gas might account for the difference in total mass and total mass in stars.

Coma is a spherical cluster of galaxies with radius 3 Mpc. The hot gas is optically thin, and we assume that it is distributed homogeneously and composed of fully ionized hydrogen. The temperature of the gas is so high that it effectively emits all its energy at x-ray wavelengths. The measured x-ray luminosity of the gas is $L_{\rm x}=5\times10^{44}$ erg sec^{-1}. We may assume $\overline{g}_{\rm III}(\nu,T)=1$.

a) Derive a formula that gives the frequency integrated emission of the gas in all directions in erg sec^{-1} cm^{-3}.

b) Derive a formula that gives the electron density as a function of $L_{\rm x}$ and other relevant quantities.

c) Calculate the total mass of the hot gas in solar mass. Recall that the proton mass $m_{\rm p}=1.66\times10^{-24}$ gr. Which fraction of the total mass of the Coma cluster resides in the hot gas component?

d) It appears that the mass in gas and stars can not account for all of the mass in the Coma cluster. What component may be responsible for the missing mass?

Exercise 8.4

The bound-free peak in figure 8.4 behaves all but hydrogen like, while it does concern hydrogen. Why is this so?

Exercise 8.5

Show that if the mass fraction of helium in a gas is Y; q_{HeII} is the ionization fraction of He II; and q_{HeIII} is the ionization fraction of He III, that the contribution of helium to the mass extinction coefficient of Thomson scattering is

$$\sigma'_{\mathrm{e}}(He) = 0.10\,Y\,(q_{\mathrm{He\,II}} + 2q_{\mathrm{He\,III}}) \tag{8.29}$$

For a gas of solar abundances $Y = 0.236$.

Exercise 8.6

Does the light of the blue sky contain spectral lines? If so, what is the origin of these spectral lines?

Exercise 8.7

Why does the sun turn red at sunset?

9

Conservation laws

In this chapter we discuss the conservation laws of mass, momentum and energy, which are faithfully fulfilled by stellar atmospheres. In essence the model atmosphere problem consists of solving the transfer problem for a suitable equilibrium between matter and radiation (see chapter 6), subject to these conservation laws. Here also, we find that in many cases one is allowed to make assumptions that greatly reduce the complexity of the problem.

The equations that essentially control the model atmosphere problem are listed in Table 9.1, together with the corresponding state parameter. We briefly highlight two of these, density and temperature. The density structure in the medium, $\rho(\mathbf{r}, t)$, is primarily determined by the conservation of mass and momentum. For relatively normal stars and planets, such as those without an outflow, it is almost always justified to assume hydrostatic equilibrium. The temperature structure, $T(\mathbf{r}, t)$, is primarily determined by the interplay between the equations of transfer and energy conservation. Often the energy in an atmosphere is transported by radiation, such that radiative equilibrium applies. In other cases it may be advantageous to transport energy by means of convection. We will not deal with this alternative means of energy transport in great detail, though we will discuss the principle of convection and the conditions in which it will occur.

9.1 Equation of state

The behavior of a gas or solid under different physical conditions is described by the equation of state (or EOS), expressing how the pressure depends on density and temperature

$$p = f(\rho, T). \tag{9.1}$$

The simplest assumption for the EOS of a gas is that it behaves as an ideal gas (see Eq. 6.28), i.e.

$$p = n\,k\,T = \frac{\rho\,k\,T}{\mu m_{\rm amu}} = \frac{\mathcal{R}\,\rho\,T}{\mu} \tag{9.2}$$

Equation	Corresponding state parameter	
Radiative transfer	Mean intensities	J_ν
	Fluxes	$\mathcal{F}_\nu$
Equation of state (EOS)	Pressure, temperature, density	$p = f(T, \rho)$
Radiative equilibrium	Temperature	T
Hydrostatic equilibrium	(Total particle) density	N, ρ
Statistical equilibrium	Populations	n_{ijk}
Charge conservation	Electron density	n_e

Table 9.1: *Summary of classical atmosphere equations and state parameters.*

where p is the gas pressure in dyn cm^{-2}, n is the particle density in cm^{-3}, ρ is the density in gr/cm^3, T is the temperature in K, and μ is the mean molecular weight of the gas in atomic mass units m_amu. The latter provides the mean mass per free particle in units of the mass of the hydrogen atom. It holds that $\rho = \mu\, m_\mathrm{amu}\, n$. The constants have their usual meaning; $\mathcal{R} = k/m_\mathrm{amu} = 8.314 \times 10^7$ erg K^{-1} mol^{-1} is the gas constant. Essentially, in an ideal gas the particles are point sources whose only interactions are perfectly elastic collisions. It is sometimes convenient when treating astrophysical problems to assume that the gas temperature and chemical composition remain constant. The former is known as the isothermal assumption, under which equation 9.2 becomes

$$\frac{p}{\rho} = \frac{k\,T}{\mu m_\mathrm{H}} = \frac{\mathcal{R}\,T}{\mu} = a^2 = \text{constant}, \tag{9.3}$$

where a is the isothermal sound speed in the gas at temperature T.

For stellar and planetary atmospheres the assumption that the gas is ideal is justified. For the interior of terrestrial planets the equation of state is set by material conditions. Seager et al. (2007, ApJ 669, 1279) summarize equations of state that assume uniform or zero temperature for a set of materials and material conditions (hence $p = f(\rho)$).

9.2 Continuity equation

The material in an astrophysical medium may not be static (i.e., $\partial/\partial t = 0$ and $v = 0$) or stationary (i.e., $\partial/\partial t = 0$), and for a general description we must take this into account. To do so, we use the basic equations of fluid mechanics. In general, there are two methods commonly used to model gas/fluid flow. One method is to use a fixed set of coordinates in space and calculate the parameters of the gas as it flows through the coordinate frame. This is known as the Eulerian method or reference frame. An alternative is to choose a set of coordinates fixed to a particle of the gas, moving with that particle, and to calculate the varying parameters in that coordinate frame (referred to as co-moving coordinates). This is

GAUSS'S DIVERGENCE THEOREM

Suppose V is a volume in three-dimensional space which is compact and has a piecewise smooth boundary S. If $\mathbf{F}$ is a continuously differentiable vector field defined on a neighborhood of V, then

$$\int_V (\nabla \cdot \mathbf{F})\, dV = \oint_S (\mathbf{F} \cdot \mathbf{dS})\,. \tag{9.4}$$

The left-hand-side is an integral over the volume V; the right-hand-side is the surface integral over the boundary of the volume V.

known as the Lagrangian method or reference frame. So, in the Eulerian reference frame think of a 1 cm^{-3} cube wire frame as a control volume being fixed in space. We monitor the rate of change of a property inside this volume, for instance density or temperature. In the Lagrangian reference frame the cube wire frame follows the motion of the flow, again we monitor the rate of change of one of its properties.

Assume an arbitrary volume V contained by a closed surface S. V and S are fixed in space, so we are here adopting the Eulerian viewpoint. The mass flowing out of V per unit time through the element of area $\mathbf{dS}$ of the surface is given by

$$\rho \mathbf{v} \cdot \mathbf{dS}, \tag{9.5}$$

where $\mathbf{v}$ is the macroscopic (i.e. systemic) velocity in cm sec^{-1} and ρ the total mass density in gr cm^{-3}, and so the net rate at which mass flows out of V through S is given by

$$\oint_S \rho \mathbf{v} \cdot \mathbf{dS} = \int_V \nabla \cdot (\rho \mathbf{v})\, dV, \tag{9.6}$$

where we have obtained the right-hand side by invoking Gauss's divergence theorem. The rate at which the mass in V decreases is given by

$$-\frac{\partial}{\partial t}\left(\int_V \rho dV\right) = \int_V \left(-\frac{\partial \rho}{\partial t}\right) dV, \tag{9.7}$$

where we can take $\partial/\partial t$ inside the integral because V is fixed in space. Obviously, the rate at which the mass in V decreases, i.e. eq. (9.7), must equal the rate at which mass flows out of V, across S, so

$$\int_V \left[\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v})\right] dV = 0, \tag{9.8}$$

and since the volume V is arbitrary, it follows that

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0 \tag{9.9}$$

everywhere. This is known as the *continuity equation*. We can expand this equation and rewrite it as

$$\frac{\partial \rho}{\partial t} + \rho\,(\nabla \cdot \mathbf{v}) + \mathbf{v} \cdot \nabla \rho = \frac{\partial \rho}{\partial t} + \rho \frac{\partial v^i}{\partial x^i} + v^i \frac{\partial \rho}{\partial x^i} = 0 \tag{9.10}$$

where the double occurrence of index i in both terms in the first right-hand-side implies a summation over the components of the coordinate system. Notice that we have implicitly assumed that there are no sources or sinks of mass in the volume V. If that would have been the case, the right hand side of Eq. 9.9 would not have been equal to zero, but would have listed the generation rate or removal rate of ρ. A positive rate would imply a mass source; a negative rate a mass sink. We briefly discuss such a possible situation later on in this section. The fact that in Eq. 9.9 no such sources or sinks exist implies that the quantity ρ is conserved.

We stress that $\partial\rho/\partial t$ is the Eulerian time derivative of the density, i.e. the rate of change of density at a fixed point in space. If we want the Lagrangian time derivative of the density, i.e. the rate of change of density moving with the fluid, we must include the contribution due to the displacement, $\mathbf{dr} = \mathbf{v}dt$, which occurs during the time interval dt. The net density change is

$$\mathrm{d}\rho = \frac{\partial \rho}{\partial t}\,\mathrm{d}t + \mathbf{dr} \cdot \nabla \rho, \tag{9.11}$$

and hence the Lagrangian time derivative of the density is

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} = \frac{\partial \rho}{\partial t} + \mathbf{v} \cdot \nabla \rho = -\rho\,(\nabla \cdot \mathbf{v})\,, \tag{9.12}$$

where the final expression is obtained by substituting from equation (9.10)[1]. The Lagrangian time derivative is called variously the 'co-moving time derivative', 'fluid-frame derivative', 'total derivative', or 'material derivative'. Though here we use the symbol d/dt for the Lagrangian derivative, it is more common to use the notation D/Dt, to distinguish it from the ordinary derivative of a function of one variable in the sense it is usually used in mathematics. We do not do that. Here d/dt is the rate of change of some characteristic (in the above case, density) of a particular element of fluid.

Mass-loss rate

For a stationary one-dimensional radial spherical flow the mass continuity equation reduces to (see section 3.1)

$$\nabla \cdot (\rho \mathbf{v}) = \frac{1}{r^2} \frac{\partial (r^2 \rho v)}{\partial r} = 0 \tag{9.13}$$

from which we find

$$4\pi r^2 \rho(r) v(r) = \text{constant} \equiv \dot{M} \tag{9.14}$$

[1] If the density of a control volume which moves with the flow remains constant, i.e. the fluid is neither compressed or expanded, the flow is said to be incompressible. It is equivalent to a vanishing divergence of the flow velocity, i.e. $\nabla \cdot \mathbf{v} = 0$.

where $\dot{M}$ is the *mass-loss rate* through a spherical surface of radius r, and v is the velocity in the radial direction. The cgs unit of mass loss is gr sec^{-1}, but in most astrophysical situations it is custom to use the unit $M_\odot$ per year: $1\ M_\odot \mathrm{yr}^{-1} = 6.303 \times 10^{25}$ gr sec^{-1}. It is straightforward to link the density of nuclei N_N (see Eq. 6.29) to the mass loss. If we introduce the *mean atomic weight*

$$\mu_\mathrm{a} \equiv \frac{1}{m_\mathrm{amu}} \sum_k m_k A_k \tag{9.15}$$

where m_k is the atomic mass of element k in atomic mass unit m_amu and A_k the abundance by number as defined by eq. (6.30), then

$$\rho(r) = m_\mathrm{amu}\, \mu_\mathrm{a}\, N_N(r) \tag{9.16}$$

such that

$$N_N(r) = \frac{\dot{M}}{4\pi m_\mathrm{amu}\, \mu_\mathrm{a}\, r^2\, v(r)} \tag{9.17}$$

Statistical equilibrium

One may also write the equation of continuity separately for each species in the medium, e.g. for Ca IV or Al X. If these species can occur in different states i we get

$$\frac{\partial n_i}{\partial t} + \nabla \cdot (n_i \mathbf{v}) = \sum_{j \neq i} (n_j P_{ji} - n_i P_{ij}) \tag{9.18}$$

Here n_i is the volume density of the specie at hand that is in energy level i and $P_{ij} = R_{ij} + C_{ij}$ the rate at which particles change from level i to j. The first term at the right-hand side of eq. (9.18) gives the number of particles per cm^3 per second that enter level i. The second term at the right-hand side gives the number of particles per cm^3 per second that leave level i.

In a static medium, or in a stationary medium in which equilibrium is reached in a time that is much shorter than the time it takes the characteristic flow velocity to change the local properties of the medium

$$\sum_{j \neq i} (n_{jk} P_{ji} - n_{ik} P_{ij}) = 0 \tag{9.19}$$

This is known as *statistical equilibrium* (SE). In general, the rate P_{ij} contains both radiative and collisional processes (see chapters 7 and 8). In NLTE it is usually assumed that the state of the gas can be described by statistical equilibrium (see section 15.5 for an example).

9.3 Momentum equation

Consider again the volume of gas V. If the gas in this volume has pressure $p(\mathbf{r}, t)$, then – pressure being the force per unit surface – the total force acting on the volume is the sum of

the external pressure on the surface. This is given by the surface integral

$$-\oint_S p\,\mathbf{dS}. \tag{9.20}$$

The minus sign naturally enters in a detailed derivation of this equation, consistently requiring the pressure (force) to act in the direction of the center of the volume. Transforming the above equation into a volume integral using Gauss's divergence theorem, the net pressure force exerted on the arbitrary volume V is

$$-\oint_S p\,\mathbf{dS} = -\int_V \nabla p\,dV, \tag{9.21}$$

and hence the net pressure force per unit volume is simply $-\nabla p$.

The equation of motion of this volume can be derived by equating the force per unit volume with the mass per unit volume multiplied by its acceleration. This is simply Newton's third law. The mass per unit volume is defined as the density ρ and the acceleration is the time derivative of the velocity, i.e. $d\mathbf{v}/dt$. So we have

$$-\nabla p = \rho\,\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t}, \tag{9.22}$$

and hence

$$\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} = -\frac{\nabla p}{\rho}. \tag{9.23}$$

Here $d\mathbf{v}/dt$ is the co-moving acceleration of the gas, so equation 9.23 is the Lagrangian formulation of the equation of motion. The Eulerian formulation is obtained by substituting for $d\mathbf{v}/dt$ using

$$\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} = \frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v}\cdot\nabla)\,\mathbf{v}. \tag{9.24}$$

For the j-th component of velocity vector $\mathbf{v}$ this implies

$$\frac{\mathrm{d}v_j}{\mathrm{d}t} = \frac{\partial v_j}{\partial t} + v^i\frac{\partial v_j}{\partial x^i}. \tag{9.25}$$

If the gas is also in a force field, for instance a gravitational field, then an extra term must be included in eq. (9.23) to account for this. The force $\mathbf{f}$ on unit volume due to an acceleration $\mathbf{g}$ is simply $\mathbf{f} = \rho\mathbf{g}$, and so eq. (9.23) becomes

$$\rho\left[\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v}\cdot\nabla)\,\mathbf{v}\right] = -\nabla p + \mathbf{f} \tag{9.26}$$

This is sometimes referred to as Euler's equation. The momentum equation primarily determines the density c.q. velocity structure of the medium.

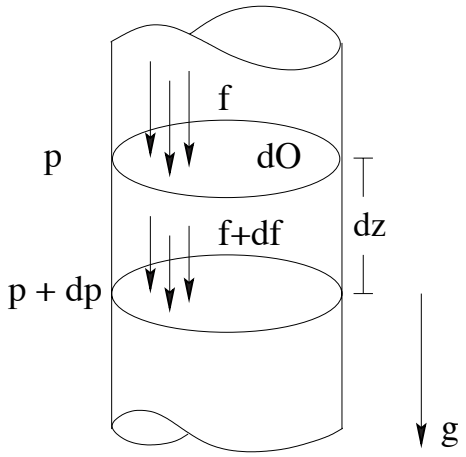


Figure 9.1: *Schematic representation of hydrostatic equilibrium. The gas that is on top exerts a force f on surface dO, causing a pressure p. The weight of the volume $dOdz$ adds an amount df to the force, such that the pressure in the interval dz increases with an amount $dp = df/dO$.*

Hydrostatic equilibrium

In a static medium (9.26) reduces to the equation of hydrostatic equilibrium

$$\nabla p = \mathbf{f} \tag{9.27}$$

Let us focus on planar atmospheres and derive the above equation again, but now in a heuristic way. In an atmosphere consisting of plane-parallel layers hydrostatic equilibrium implies an equilibrium between the force generated by the gradient of the gas pressure and the gravitational force at the surface, which is assumed constant

$$g = \frac{GM}{R^2} \tag{9.28}$$

where G is the gravitational constant. The gravitational force may be considered constant if *i)* the thickness of the atmosphere is much less than the radius R – the inherent assumption of a planar atmosphere, and *ii)* the mass of the atmosphere is much smaller than the total mass M, which holds for all stars and planets.

Consider an elementary volume, with cross-section dO and length dz in the direction normal to the planar layers (see figure 9.1). The difference in pressure between the top and bottom of the volume is the force per unit surface of the mass of the volume. This force df is the product of density ρ, volume, and gravity g. For the pressure difference we find $dp = \mathrm{d}f/dO = -\rho g_\star \, dz$, therefore

$$\frac{\mathrm{d}p}{\mathrm{d}z} = -\rho g \tag{9.29}$$

We now assume that the total pressure in the atmosphere is dominated by the gas pressure, and that radiation pressure can be neglected. This allows us to use the ideal gas law eq. (9.2).

If we further assume that the atmosphere is isothermal, we find after substitution of our EOS in (9.29) for the solution of this first order differential equation

$$\rho(z) = \rho(z_\circ) \exp\left[-\left(z - z_\circ\right)/H\right], \tag{9.30}$$

where

$$H = \frac{\mathcal{R}T}{\mu g} = \frac{a^2}{g} \tag{9.31}$$

is the *density scale height* of the isothermal atmosphere and $z_\circ$ is an arbitrary reference point along the z-axis. H measures the distance over which the density changes with a factor e^{-1}. For our sun this is about 150 km, a sizable fraction of the total geometrical thickness of the photosphere ($\sim$ 500 km) though only a tiny fraction of the solar radius (700 000 km). For the M1.5 Iab supergiant Antares (α Sco) with $T_{\rm eff} = 3600$ K, $R_\star = 650 R_\odot$ and $\log g = 0$ one finds $H = 5\,000\,000$ km. This too is a sizable fraction of the thickness of the photosphere of Antares. Taking $T = 250$ K, $\mu = 29$ and $g = 981\,\mathrm{cm\,s^{-2}}$, we find for Earth's atmosphere $H = 7.3$ km. Some more details about our planet's atmosphere are given in Sect. 12.4. For Jupiter, assuming a typical atmospheric temperature of 124 K, $\log g = 3.39$, and $\mu = 4.3$, yields $H = 10$ km. Again, a sizeable fraction of the total extent of Jupiter's atmosphere. For hot Jupiter-like exoplanets, the density scale height can be several hundred kilometers. Notice that the expression for H implies that for an increasing gravity the atmosphere will become more compact.

Radiation pressure

For extremely luminous stars (such as O-type stars, LBVs and Wolf-Rayet stars) the pressure that is exerted by the photons $p_{\rm R}$ (see eq. 3.36) is an important component of the total pressure $p = p_{\rm G} + p_{\rm R}$. Note that, so as to avoid confusion, we have introduced the subscript G to identify the gas pressure. We may rewrite the equation of hydrostatic equilibrium into

$$\frac{{\rm d}p_{\rm G}}{{\rm d}m} = {\rm g}_\star - \frac{{\rm d}p_{\rm R}}{{\rm d}m} = {\rm g}_\star - {\rm g}_{\rm R}(m) \tag{9.32}$$

where ${\rm g}_{\rm R}(m)$ is the radiative acceleration and m is the *column mass* (dimensions gr cm^{-2}). The latter serves as an independent measure of the depth in the atmosphere and is measured from the outside in, i.e.

$$dm \equiv -\rho\, dz \tag{9.33}$$

such that at depth $z_\circ$

$$m(z_\circ) = \int_{z_\circ}^{\infty} \rho(z) dz \tag{9.34}$$

gives the total mass of a column of cross-section unit cm^2 between the observer and z. The column mass is a very favourable variable as it allows for a very simple solution of the pressure structure in the atmosphere, $p(m) = {\rm g}_\star m + \text{constant}$.

Using eq. (3.34), (3.36), (4.35), and (4.44) we find for the outward directed acceleration due to the gradient of the radiation pressure

$$\mathrm{g_R}(m) \equiv \frac{\mathrm{d}p_\mathrm{R}}{\mathrm{d}m} = \frac{4\pi}{c}\int_0^\infty \chi'_\nu H_\nu \, d\nu \tag{9.35}$$

As the accelerations due to gravity $\mathrm{g}_\star$ and radiation $\mathrm{g_R(m)}$ have opposite signs one can have a situation where the net force in the atmosphere is zero (or even outward directed): the gas at this location is no longer bound to the star but "floats" on the photons (loosely speaking). This situation, for which $\mathrm{g}_\star = \mathrm{g_R}$, is known as the *Eddington limit*. In hot O stars, where hydrogen is fully ionized throughout the atmosphere, the extinction is dominated by Thomson scattering (see § 8.4). Using the fact that this process is wavelength independent, and using eq. (3.22), (3.23) and (3.25) we may derive an expression for the luminosity that is required to reach this situation of no net force in the atmosphere. We find for this *Eddington luminosity*

$$L_\mathrm{E} = \frac{4\pi c G M_\star}{\sigma'_\mathrm{e}} = 1.3\times10^4 \frac{1}{\sigma'_\mathrm{e}} \frac{M_\star}{M_\odot} \quad L_\odot \tag{9.36}$$

It is thought that the Eddington limit plays an important role in the life of very massive stars (with initial masses $M \gtrsim 40 M_\odot$) when they reach an evolutionairy phase in which they are spectroscopically identified as LBV (see § 2.2).

Atmospheres that are not in hydrostatic equilibrium

In the above discussion we have assumed that the forces due to pressure gradients and gravity are in equilibrium. Yet there are stars that have, for a range of reasons, a stationary stellar wind. This wind causes the atmosphere to be extended. In many cases the stellar wind does not have a significant influence on the continuum emission of the star. For luminous O, LBV and Wolf-Rayet stars this, however, *is* the case (see § 2.2). Notably Wolf-Rayet stars may develop stellar winds that are so strong that the continuum radiation is formed in the wind. Acoustic waves in red giants and supergiants, and pulsations in asymptotic giant branch stars can also lead to very extended atmospheres. Stellar winds of hot stars will be discussed in more detail in chapter 17.

9.4 Energy equation

Conservation of energy is expressed by the *energy equation*

$$\frac{\partial}{\partial t}\left(\frac{1}{2}\rho v^2 + \rho\epsilon\right) + \nabla\cdot\left[\left(\frac{1}{2}\rho v^2 + \rho\epsilon\right)\mathbf{v}\right] + \nabla\cdot(p\mathbf{v}) = \mathbf{f}\cdot\mathbf{v} - \nabla\cdot(\mathcal{F}_\mathrm{rad} + \mathcal{F}_\mathrm{conv} + \mathcal{F}_\mathrm{con}) \tag{9.37}$$

ϵ is the specific (i.e. per unit mass) internal energy and $v^2/2$ the specific bulk kinetic energy. Think of ϵ as the grand total of microscopic manifestations of kinetic energy (e.g. translation, rotation, vibration) and potential energy (e.g. excitation, ionisation, chemical bonds) of the system's particles. The first term describes the change of energy in time. The second term specifies the change in energy because of movement through the medium. $\nabla \cdot (p\mathbf{v})$ is the work done by the gas pressure p to accomplish a change in the specific volume (so the change of volume per unit mass, i.e. per $1/\rho$). $\mathbf{f} \cdot \mathbf{v}$ is the work done by external forces. Heating by radiation, convection and conduction are described by the divergences of $\mathcal{F}_{\rm rad}$, $\mathcal{F}_{\rm conv}$ and $\mathcal{F}_{\rm con}$, i.e. the total radiative, convective and conductive flux, respectively. Heat conduction occurs as rapidly moving or vibrating atoms and molecules interact with neighboring atoms and molecules, transferring some of their energy to these neighboring particles. In other words heat is transferred by conduction when adjacent atoms vibrate against one another, or as electrons move from one atom to another. Conduction is the most significant means of heat transfer within a solid or between solid objects in thermal contact. Fluids and especially gases are less conductive. Convection is the transfer of heat from one place to another by the movement of fluids or gases. Radiation is, as we have discussed, a means to transport energy by means of photons. In principle other terms could be added to the energy equation, such as terms due to viscous processes. The energy equation primarily determines the temperature structure of the medium.

Radiative equilibrium

In a static medium in which all energy is transported in the form of radiation the energy equation reduces to

$$\nabla \cdot \mathcal{F}_{\rm rad} = 0 \tag{9.38}$$

This is the equation of *radiative equilibrium*. For a planar atmosphere we find $d\mathcal{F}/dz = 0$, where we have dropped the subscript "rad" as hereafter we will only speak of radiative flux. So, the same amount of total flux passes through each planar layer. This is to be expected. As no energy is generated in the atmosphere of a normal star (for instance by thermonuclear reactions) or planet the only thing that needs to be done is to transport the energy outward. The constant total flux is given by eq. (3.23), which we here repeat

$$\mathcal{F} = \text{constant} = \sigma T_{\rm eff}^4 \tag{9.39}$$

Note that the effective temperature $T_{\rm eff}$ does *not* represent a physical temperature, but is a measure of the total flux (see also § 6.6).

For a spherical atmosphere eq. (9.38) implies that $\partial(r^2\mathcal{F})/\partial r = 0$, or

$$4\pi r^2 \mathcal{F}(r) = \text{constant} = L \tag{9.40}$$

where L is the luminosity of the star. So, for the case of spherical shells the total flux, and therefore the effective temperature, depends on distance. To conserve the concept of effective temperature (as best as possible) it is custom to identify $T_{\rm eff}$ with the total flux at the stellar surface $R_\star$ (see eq. 3.25).

Equations 9.39 and 9.40 are *global* representations of the condition of radiative equilibrium. The total flux that enters the "inner boundary" of the atmosphere should also leave at the "outer boundary". We can however also formulate the requirement of radiative equilibrium in a *local* representation using the zero-order moment of the equation of transfer (eq. 4.43 or 4.46). For a planar geometry, after substituting eq. 4.35, we obtain

$$\frac{\mathrm{d}H_\nu}{\mathrm{d}z} = \chi_\nu \left(S_\nu - J_\nu\right). \tag{9.41}$$

Integrating over frequency yields $dH/dz = 1/4\pi \cdot d\mathcal{F}/dz$, which equals zero. Multiplying by 4π yields

$$4\pi \int_0^\infty \chi_\nu(z) \left[S_\nu(z) - J_\nu(z)\right] \, d\nu = 0. \tag{9.42}$$

An identical equation is found for the case of spherical shells, with z is replaced by r. Each volume element of gas must fulfill the above requirement. This is why this is called the local representation of radiative equilibrium. The equation describes that the total amount of energy that is absorbed by an elementary volume of gas per second ($4\pi \int_0^\infty \chi_\nu J_\nu \, d\nu$) must be equal to the total amount of energy that is emitted from the volume in the same time interval ($4\pi \int_0^\infty \chi_\nu S_\nu \, d\nu = 4\pi \int_0^\infty \eta_\nu \, d\nu$).

The effect of scattering on radiative equilibrium

In eq. (9.42) we may replace the total extinction coefficient χ_ν by the "true" absorption coefficient κ_ν (see eq. 4.4) and the source function by the local Planck function. The reason for this is that scattering contributions drop out of the equation. Obviously, this is what one would intuitively expect should happen: the scattering of a photon does not lead to a change in the energy balance in the elementary volume and therefore can not contribute to eq. (9.42). To illustrate this we use a proto-type source function (see also eq. 4.33) that contains both a thermal and a scattering component, i.e.

$$S_\nu = \frac{\kappa_\nu B_\nu + \sigma_\nu J_\nu}{\kappa_\nu + \sigma_\nu} \tag{9.43}$$

Let us first briefly discuss this equation. Thermal emission, as expected, is described by the Kirchhoff-Planck relation (6.3), i.e. $\eta_\nu^{\mathrm{th}} = \kappa_\nu B_\nu$. The contribution of scattering processes to the emission is given by $\eta_\nu^{\mathrm{sc}} = \sigma_\nu J_\nu$, where we have assumed that the scattering process is isotropic and coherent (for coherent scattering, see § 8.4). After all, for isotropic scattering the fraction of energy that is scattered out of a beam of light that is incident from direction $\mathbf{n}$, into a beam of solid angle $d\omega$ in for instance the direction of the observer, is $d\omega/4\pi$. For coherent scattering the contribution to the emission into this direction is $d\eta_\nu^{\mathrm{sc}}(\mathbf{r}) = \sigma_\nu(\mathbf{r}) I_\nu(\mathbf{r}, \mathbf{n}) d\omega/4\pi$. Integrating over all incident directions yields $\eta_\nu^{\mathrm{sc}} = \sigma_\nu J_\nu$.

Substitution of the proto-type source function in eq. (9.42) results in

$$4\pi \int_0^\infty \kappa_\nu(z) \left[B_\nu(z) - J_\nu(z)\right] \, d\nu = 0 \tag{9.44}$$

This equation shows that the total thermal emission $\int_0^\infty \kappa_\nu B_\nu \, d\nu$, which fixes the *local* temperature T, is determined by the mean intensity. The value of J_ν depends upon the *global* properties of the atmosphere because it follows from a solution of the transfer equation. Thus the temperature at a given point in the atmosphere is to some extent determined by the temperature at all other points and, at the same time, helps to establish the temperature structure elsewhere. This nonlocalness in the problem is a result of radiative transfer, by which photons moving from one point in the medium to another give rise to a fundamental coupling (i.e. interdependence) of the properties at those points.

9.5 Convection

In a stellar atmosphere energy can also be transported by convection, i.e. by large scale motions of gas elements that obtain an excess energy content at some place and that deposit this excess energy at some other place. Convection becomes important in the somewhat deeper layers of the atmospheres of late type stars, starting at about middle F-stars. In the outermost part of the atmosphere, where $\tau < 1$ and photons can freely escape, the transport of energy by radiation is always more efficient than by convective motions. For stars of earlier spectral type than mid-F all of the atmosphere is in radiative equilibrium.

At the solar surface the convective motions are visible as a complex granulous pattern (see figure 9.2). The rising convective cells, called *granules*, are about 100 K warmer than the surrounding relatively dark cells. Because of this temperature difference the rising granules are about 25 percent brighter. The velocity of the cells is between 1 and 2 km sec^{-1}. The cells that go down have a velocity that typically is twice as large. There are at least two reasons why the falling gas moves faster than the rising gas. First, the darker areas cover less than half of the solar surface. Conservation of mass flow thus requires the downward velocities to be larger than the upward velocities. More important, the continuum extinction is lower in the cooler areas, such that we see deeper layers, where the convective velocities are larger.

Convective motions are turbulent and consist of a complicated hierarchy of "eddies" or "bubbles" moving and interacting in an extremely involved way. Turbulence is a complex mathematical problem for which a definitive theory does not yet exist. Instead one adopts a phenomenological description that includes free scaling parameters. In this *mixing length theory* it is assumed that the convective cells travel a certain distance (the mixing length) before they deposit their excess thermal energy into their surroundings. We will not give a detailed description of the physics of energy transfer by convection, rather we focus on deriving a criterion that will allow us to determine whether energy is transported by radiation or by convection.

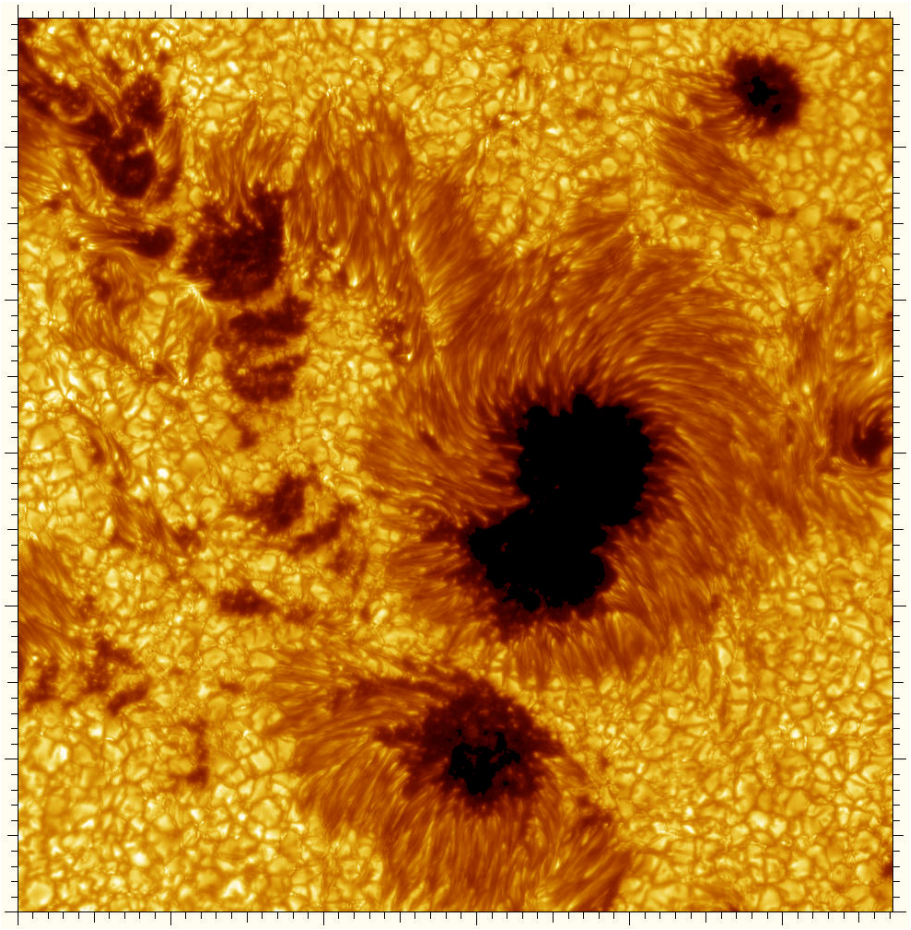


Figure 9.2: *Detailed image of the surface of the sun. The tick marks are 1000 km apart. Though what we see here is part of a sunspot group near disk center, the irregular granular pattern can clearly be seen around the sunspot structures. The granular pattern is caused by temperature variations over the surface. The individual cells, or* granules *are the signature of convective motions and exist for about five minutes. The rising granules are about 100 K warmer compared to the relatively dark inter-granular regions that are falling; this higher T corresponds to a brightness difference of about 25 percent. The velocity of the rising granules is about 1 to 2 km sec^{-1}; the velocities of the falling material is about twice as large. From: The institute for Solar Physics, Sweden.*

Schwarzschild instability criterion

Let us forget about the complex of turbulent gas motions, and focus on the movement of a single identifiable "bubble" of gas. We assume this gaselement is hotter than its surroundings. In order for the bubble *not* to radiate this surplus of thermal energy the total optical depth measured over the dimensions of the bubble should be large, i.e. $\tau > 1$. We also assume that the motion of the bubble is subsonic, such that it will be in pressure equilibrium with its surroudings (an equilibrium that sets itself with the velocity of sound). Because of this the

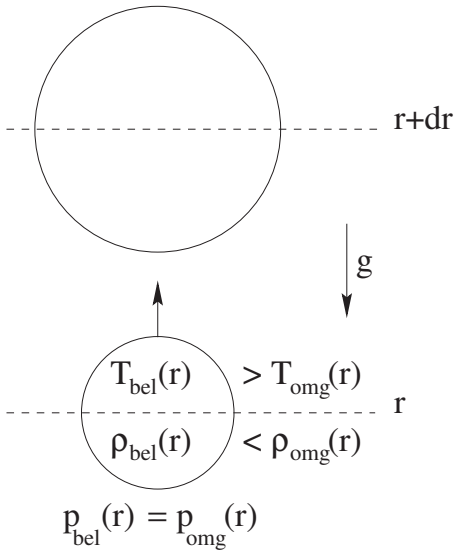


Figure 9.3: *Schematic representation of convection. A gas bubble having a temperature slightly above that of its surrounding, and that does not exchange heat with these surroundings, will rise because the density in the bubble is lower than that of its surroundings.*

density in the bubble will be lower than that of its surroundings. According to Archimedes law the bubble will rise (see figure 9.3). We further assume that all during the rise of the bubble non of the surplus in energy will leak into the ambient medium – i.e. the convective bubble behaves adiabatically. The change in density in the cel, after it has ascended a distance dr, can then be described by $(d\rho)_{\rm cel} = (d\rho/dr)_{\rm ad}\, dr$. In a radiative surroundings the change in density over the same distance will be $(d\rho)_{\rm sur} = (d\rho/dr)_{\rm rad}\, dr$. As long as the density in the bubble is lower than that of its surroundings it will continue its rise upward in the atmosphere, and convection will persist. Convection thus occurs when $(d\rho/dr)_{\rm cel} < (d\rho/dr)_{\rm sur}$ (Recall that both density gradients are negative), or, phrased differently

$$\left.\frac{\mathrm{d}\log\rho}{\mathrm{d}\log p}\right|_{\rm ad} > \left.\frac{\mathrm{d}\log\rho}{\mathrm{d}\log p}\right|_{\rm rad} \tag{9.45}$$

(the sign flips as also dp/dr is negative). For adiabatic circumstances *Poisson's law* is valid

$$p = \text{constant}\,\rho^{\gamma} \tag{9.46}$$

where the *adiabatic index* γ is the ratio between the specific heat at constant pressure and constant volume. One gets

$$\left.\frac{\mathrm{d}\log\rho}{\mathrm{d}\log p}\right|_{\rm ad} = \frac{1}{\gamma} \tag{9.47}$$

Using the ideal gas law eq. (9.2) we find for this gradient in a radiative environment

$$\left.\frac{\mathrm{d}\log\rho}{\mathrm{d}\log p}\right|_{\rm rad} = 1 - \frac{\mathrm{d}\log T}{\mathrm{d}\log p} + \frac{\mathrm{d}\log\mu}{\mathrm{d}\log p} \tag{9.48}$$

Substituting the results (9.47) and (9.48) into the convection criterion (9.45) yields

$$\left.\frac{\mathrm{d}\log T}{\mathrm{d}\log p}\right|_{\mathrm{rad}} > \frac{\gamma-1}{\gamma} + \frac{\mathrm{d}\log\mu}{\mathrm{d}\log p} = \left.\frac{\mathrm{d}\log T}{\mathrm{d}\log p}\right|_{\mathrm{ad}} \tag{9.49}$$

The last equality on the right-hand side can be verified by substitution of the ideal gas law into eq. (9.46) and subsequent differentiation. The above equation is known as the *Schwarzschild instability criterion*, after the person that thought of it. In terms of the temperature gradient dT/dr we may write the criterion for convection to occur as

$$\left|\frac{\mathrm{d}T}{\mathrm{d}r}\right|_{\mathrm{rad}} > \left|\frac{\mathrm{d}T}{\mathrm{d}r}\right|_{\mathrm{ad}} \tag{9.50}$$

This shows that an atmosphere will opt for the smallest temperature gradient: is $|dT/dr|_{\mathrm{rad}}$ larger than $|dT/dr|_{\mathrm{ad}}$ it will switch to the transport of energy by convection.

Condition for convection to occur

Using the diffusion approximation 4.68, which for the radiative temperature gradient yields

$$\frac{\mathrm{d}T}{\mathrm{d}z} = -\frac{3}{16\sigma}\frac{\chi\prime_{\mathrm{R}}\rho\mathcal{F}}{T^3} \tag{9.51}$$

where $\chi\prime_{\mathrm{R}}$ is the Rosseland mean mass extinction coefficient (dimensions cm^2 gr^{-1}), i.e. $\chi_{\mathrm{R}} = \chi\prime_{\mathrm{R}}\rho$, and equation (9.29) describing hydrostatic equilibrium, we may obtain an explicit formulation of the radiative gradient (9.49). We find

$$\left.\frac{\mathrm{d}\log T}{\mathrm{d}\log p}\right|_{\mathrm{rad}} = -\frac{3}{16\sigma}\frac{\chi\prime_{\mathrm{R}}\mathcal{F}p}{\mathrm{g}_\star T^4} = -\frac{3}{16}\frac{\chi\prime_{\mathrm{R}}p}{\mathrm{g}_\star}\left(\frac{T_{\mathrm{eff}}}{T}\right)^4 \tag{9.52}$$

The atmosphere will thus become unstable for convection if the extinction becomes too large. For increasing extinction it will be harder to transport energy by means of photons, since radiation gets effectively trapped. The star has two ways to solve for this problem. One way is to increase in size, as $(d\log T/d\log p)_{\mathrm{rad}} \propto p/T^4 \propto \rho/T^3$, and therefore the value of the radiative gradient can go down if at the position of temperature T the density would be lower. In other words, the star needs to swell. The second possibility is for the star to switch to a transport of energy by convection the moment that criterion (9.49) is fulfilled.

It is more likely for convection to take over if the adiabatic gradient is small. This may happen when changes in the ionization or in the chemistry of the gas occur. We first discuss the effects of chemistry. Stars of relatively late spectral type (G or later) show a strong increase in molecular abundances. The formation of molecules changes the adiabatic index γ from 5/3 for a fully mono-atomic gas toward 7/5 for a fully di-atomic gas. The result is that the factor $(\gamma-1)/\gamma$ in eq. (9.49) goes down from 0.4 to 0.29. The presence of molecules thus increases the chance for convection to occur. The adiabatic gradient is also lower in regions where hydrogen (or helium) gets ionized. Here the gradient $d\log\mu/d\log p$ is negative, lowering the

threshold at which convection starts. Progressing from deep layers towards the surface, i.e. in the direction of decreasing temperature, the extinction of hydrogen will increase strongly in the zone in which this element recombines, notably due to bound-free processes. This requires a steeper temperature gradient for energy to be transported by radiation, and thus increases the likelihood that convection takes over. The effect of recombination in a convective cel (that does not exchange heat with its environment) is that extra energy is added to the gas (13.6 eV per recombination of electron and proton), causing the bubble to expand further, which sustains convection. Zones of molecule formation in extremely cool stars cause a similar effect. Such zones are therefore advantageous for two reasons, both for initiating and sustaining convection.

Exercise 9.1

a) Start with eq. (9.10) and show that in planar geometry the time-dependent continuity equation is given by

$$\frac{1}{\rho}\frac{\partial \rho}{\partial t} + \frac{\partial v_z}{\partial z} + \frac{v_z}{\rho}\frac{\partial \rho}{\partial z} = 0 \tag{9.53}$$

b) Start with eq. (9.10) and show that in spherical geometry the time-dependent continuity equation for radial flow is given by

$$\frac{1}{\rho}\frac{\partial \rho}{\partial t} + v\left[\frac{1}{v}\frac{\partial v}{\partial r} + \frac{2}{r} + \frac{1}{\rho}\frac{\partial \rho}{\partial r}\right] = 0, \tag{9.54}$$

where v is the velocity in the radial direction.

c) Start with eq. (9.14) and show that for a stationary radial flow the continuity equation is given by

$$\frac{1}{v}\frac{\partial v}{\partial r} + \frac{2}{r} + \frac{1}{\rho}\frac{\partial \rho}{\partial r} = 0 \tag{9.55}$$

Exercise 9.2

The total number of particles of species k is given by

$$N_k = \sum_i n_{ik} \tag{9.56}$$

If the mass of the individual particles of species k is m_k, then

$$\rho = \sum_k m_k N_k \tag{9.57}$$

a) Give the continuity equation of species k

b) Show that from a summation over all species one recovers the continuity equation (9.9) for the mass density.

Exercise 9.3

a) Assume that in a planar geometry the only force field present is f_z and that this field is acting in the z-direction only. Start with eq. (9.26) and show that in this geometry the time-dependent momentum equation is given by

$$\rho\left[\frac{\partial v_z}{\partial t} + v_z\frac{\partial v_z}{\partial z}\right] = -\frac{\partial p}{\partial z} + f_z \tag{9.58}$$

b) Assume that in a spherical geometry the only force field present is f_r and that this field is acting along the radial direction only. Start with eq. (9.26) and show that in this geometry the time-dependent momentum equation for radial flow is given by

$$\rho \left[\frac{\partial v}{\partial t} + v \frac{\partial v}{\partial r} \right] = -\frac{\partial p}{\partial r} + f_r \tag{9.59}$$

Exercise 9.4

For general equations of state, the speed of sound is given by

$$v_s^2 = \frac{\partial p}{\partial \rho} \tag{9.60}$$

where the derivative is taken with respect to adiabatic change. It follows from eq. (9.2) that for an isothermal medium and a constant mean molecular weight μ, $v_s = \sqrt{\mathcal{R}T/\mu}$. Assume that the only force field present is gravity, such that $f_r = \rho\, g$ with $g = -GM/r^2$. Use the results eq. (9.55) and eq. (9.59) and show that for an isothermal medium in which the mean molecular weight is constant, the equation of motion for a stationary radial flow can be written as

$$\frac{1}{v}\frac{\partial v}{\partial r} = \left[\frac{2v_s^2}{r} - \frac{GM}{r^2} \right] / \left[v^2 - v_s^2 \right]. \tag{9.61}$$

This equation has a singularity at the point where $v(r) = v_s$, where the denominator is zero and $\partial v/\partial r$ becomes infinite (which cannot be physical), unless the numerator is also zero there. As this equation traces the change of velocity in only one direction, i.e. r, the partial derivative $\partial/\partial r$ may be replaced by the total derivative d/dr. We return to this equation in Chapter 17.

Exercise 9.5

a) Compute the density scale height of the Earth atmosphere at the surface. Assume our atmosphere to be isothermal.

b) Compute the gas density at the Earth surface knowing that the surface pressure is 1.013×10^6 barye (cgs unit) or 1.013×10^5 Pa (SI unit).

c) Compute the total column mass m (Eq. 9.34) and total gas mass of Earth's atmosphere.

Exercise 9.6

We are concerned with the equation of hydrostatic equilibrium for a static, spherical medium,

$$\frac{\mathrm{d}p}{\mathrm{d}r} = -\rho g = -\rho \frac{GM_\star}{r^2} = -\rho \frac{GM_\star}{r_\star^2} \left(\frac{r_\star}{r}\right)^2 = -\rho g_\star \left(\frac{r_\star}{r}\right)^2 , \tag{9.62}$$

where $r_\star$ is a reference depth (think of it as being the stellar radius) and $g_\star$ the acceleration of gravity at the reference radius.

a) Show that for the isothermal case, i.e. $T(r) = T_\star$

$$\rho(r) = \rho(r_\star) \exp\left[-\frac{(r - r_\star)}{H_\star} \frac{r_\star}{r}\right] , \tag{9.63}$$

where $H_\star = \mathcal{R}T_\star/\mu g_\star$ is the scale height at the reference depth.

b) Show that for a power-law temperature structure

$$T(r) = T_\star \left(\frac{r}{r_\star}\right)^{-\beta} , \tag{9.64}$$

with β the temperature profile exponent,

$$\rho(r) = \rho(r_\star) \left(\frac{r}{r_\star}\right)^{\beta} \exp\left\{\frac{-r_\star}{H_\star(\beta - 1)} \left[\left(\frac{r}{r_\star}\right)^{\beta-1} - 1\right]\right\} \tag{9.65}$$

Exercise 9.7

a) Derive that the acceleration due to the gradient in the radiation pressure, $\mathrm{g_R}$, is given by equation (9.35).

b) Show that if $\mathrm{g}_\star = \mathrm{g_R}$, the Eddington luminosity is given by equation (9.36), assuming the continuum extinction is dominated by Thomson scattering.

Exercise 9.8

This could be a nice exam question. Consider a homogeneous isothermal planar atmosphere, with temperature $T = 6\,000\,\mathrm{K}$; mean molecular weight $\mu = 0.6$, and a gravity $\log g = 4.4$ in c.g.s. units. The continuum mass extinction coefficient $\chi'_{\mathrm{cont}} = 0.6\,\mathrm{cm^2\,gr^{-1}}$. The atmosphere is in hydrostatic equilibrium. The gas may be considered to be ideal. The gas constant $\mathcal{R} = 8.31 \times 10^7$ (cgs).

a) Compute the density scale height in this atmosphere?

b) Derive an expression for the optical depth at position $z_\circ$ in the atmosphere when it is given that $\rho(z_\circ) = \rho_\circ$.

c) Compute the density in gr cm^{-3} at $\tau_{\rm cont} = 1$.

d) Explain why the (pressure broadened) hydrogen lines in supergiants are so narrow (flip ahead to Section 13.2 for more information).

Exercise 9.9

Show that the local constraint of radiative equilibrium, eq. (9.42), is already contained in eq. (4.30) and (4.31).

10

Grey, planar, LTE atmosphere in hydrostatic & radiative equilibrium

In this chapter, we combine all that we have discussed so far and construct a model for a planar, LTE, grey atmosphere in hydrostatic and radiative equilibrium. To be able to do this completely analytical, we apply the Eddington approximation. In this way, we can derive expressions for the temperature structure $T(z)$ and density structure $\rho(z)$. These two properties are the essence of the model atmosphere.

The definition of a *grey medium* is that the extinction coefficient is independent of frequency, i.e. $\chi_\nu = \chi$. In many cases this is a very unrealistic assumption. For instance, think about absorption lines and ionization edges. These show an extreme frequency dependent behavior. The only real example of a grey extinction is that of Thomson scattering on free electrons. Still, studying the grey atmosphere problem is extremely useful. First, because the grey problem very nicely illustrates how the temperature structure is the result of a combined solution of the transfer equation and the energy conservation equation. Second, because the end result for $T(r)$ compares remarkably well to those obtained from advanced models, and third, because the latter implies that this simple model may serve as an excellent starting model for the iterative methods that are required to construct LTE and NLTE atmosphere models.

10.1 Description of the grey atmosphere

We assume that the atmosphere consists of planar layers. The equation of transfer for a grey medium follows from eq. (4.36) and is given by

$$\mu \frac{\mathrm{d}I_\nu}{\mathrm{d}\tau} = I_\nu - S_\nu \tag{10.1}$$

Note that the optical depth is no longer frequency dependent, as $d\tau_\nu = -\chi_\nu\, dz = -\chi\, dz = d\tau$. Integration of this equation over frequency yields

$$\mu \frac{\mathrm{d}I}{\mathrm{d}\tau} = I - S \tag{10.2}$$

where I is the total specific intensity (eq. 3.3) and S the total source function (eq. 4.34).

We require the atmosphere to be in radiative equilibrium. From eq. (9.42) it follows that

$$\int_0^\infty \chi S_\nu \, d\nu = \int_0^\infty \chi J_\nu \, d\nu \tag{10.3}$$

Because $\chi \neq \chi_\nu$, the equation reduces to $S = J$. If the source function can be written as eq. (9.43) it also holds that $J = B$ (see eq. 9.44), and therefore $S = B$. Using Stefan-Boltzmann's law (eq. 6.11) we finally obtain $J = (\sigma/\pi)T^4$. This implies that if we derive a solution for $J(\tau)$, we may immediately couple it to the temperature, automatically fulfilling the requirement of energy conservation.

Integration of the 0^{th} order moment of the grey equation of transfer over frequency (eq. 4.43) yields

$$\frac{\mathrm{d}H}{\mathrm{d}\tau} = J - S = 0 \tag{10.4}$$

implying that the total Eddington flux H = constant = $\sigma\, T_{\mathrm{eff}}^4/4\pi$.

Integration of the 1^{st} order moment of the grey equation of transfer over frequency (eq. 4.44) results in

$$\frac{\mathrm{d}K}{\mathrm{d}\tau} = H \tag{10.5}$$

which has an elementary solution $K(\tau) = H \times (\tau + \mathrm{constant})$[1]. If we also integrate the Eddington factor f_ν over frequency, which gives $f = K/J$, we find

$$3f(\tau)J(\tau) = 3H(\tau + \mathrm{constant}) \tag{10.6}$$

and for the temperature as a function of optical depth

$$3f(\tau)T^4(\tau) = \frac{3}{4}\, T_{\mathrm{eff}}^4(\tau + \mathrm{constant}) \tag{10.7}$$

To know the exact behavior $T(r)$ we thus need to determine the run of $f(\tau)$ throughout the atmosphere. As we have seen in § 3.6, deep in the atmosphere, where the radiation field is almost isotropic, $f \to 1/3$, such that

$$J(\tau) \to 3H\tau \qquad\qquad \forall\ \tau \gg 1 \tag{10.8}$$

At depth, the mean intensity is a linear function of optical depth. Close to the surface we expect that $f(\tau) \neq 1/3$, and thus $J(\tau)$ can not be described by this simple relation. To account

[1] Do not get confused: the solution is $K(\tau) = H\tau + \mathrm{constant}\prime$. As H is a constant one may define $\mathrm{constant}\prime = H \times \mathrm{constant}$.

for this, we could define a general relation for the exact expression of the mean intensity that looks like

$$J(\tau) = 3H\,(\tau + q(\tau)) \tag{10.9}$$

where the anisotropy of the radiation field has been absorbed in the function $q(\tau)$. This $q(\tau)$ is known as the *Hopf-function*. There is an elegant, though somewhat complex analytical solution of the Hopf-function that requires numerical integration. We will return to this solution below. Using the Hopf-function we find for the general solution of the grey temperature structure

$$T^4(\tau) = \frac{3}{4}\,T_{\rm eff}^4\,(\tau + q(\tau)) \tag{10.10}$$

Eddington approximation

To get a fully explicit solution of the grey temperature structure we pursue an idea of Eddington and assume that everywhere in the atmosphere $f = 1/3$. This assumption, you'll never guess, is called the *Eddington approximation*. For the total mean intensity we obtain from eq. (10.6)

$$J(\tau) = 3H(\tau + \text{constant}) \tag{10.11}$$

This is a linear function of optical depth, which therefore also holds for the source function, as in a grey medium $J = S$. It implies that also the Eddington-Barbier approximation is valid (see § 4.6). From an integration of eq. (4.56) over frequency we know that the emerging total flux is given by

$$H^+(0) = H = \frac{1}{4}S(\tau = 2/3) \tag{10.12}$$

Combining the above two equations one finds for the constant a value of 2/3. For the temperature we find

$$T^4(\tau) = \frac{3}{4}\,T_{\rm eff}^4(\tau + \frac{2}{3}) \tag{10.13}$$

In the Eddington approximation, therefore, $q(\tau) \equiv 2/3$. The real value of the Hopf function smoothly varies between $q(0) = 0.577$ and $q(\infty) = 0.710$, which is never that far off from the Eddington approximation. Table 10.1 gives for a number of optical depths the exact solution of the Hopf function.

Result 10.13 shows that in a grey atmosphere the temperature structure follows from the equation of transfer and the energy equation. The momentum equation, let us assume hydrostatic equilibrium, does not enter in the problem. In other words, the temperature in a grey atmosphere, as a function of optical depth, does not depend on the gravity $\mathrm{g}_\star$. However, the equation of hydrostatic equilibrium does control the relation between optical depth and geometrical depth. We will return to this below.

The grey temperature structure is a monotonous function of optical depth. Why is this so? Actually, it follows from the condition of radiative equilibrium, which requires that the total radiative flux is constant throughout the atmosphere. The relation between the radiative flux and the mean intensity measures the degree of anisotropy of the medium (see § 3.6). With

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

Table 10.1: *The exact solution of the Hopf function $q(\tau)$ for a planar atmosphere. The limit value $q(0) = 1/\sqrt{3}$.*

increasing (optical) depth in the atmosphere the degree of anisotropy decreases. The only means of keeping the flux constant is by increasing the energy density of radiation (which is proportional to J). As $J \propto T^4$, the temperature has to increase.

The Eddington approximation also predicts that $T = T_{\rm eff}$ for $\tau = 2/3$. This result confirms that the "effective depth" of continuum formation is at $\tau \sim 2/3$ (see also eq. 4.56). Note that a photon that is emitted at position $\tau = 2/3$ in the outward direction has a chance of order $e^{-2/3} \sim 0.5$ to escape from the atmosphere by direct flight. This intuitively agrees with what one would expect about the place where the continuum is formed. Finally: equation (10.13) predicts that the temperature at the outer boundary $T(0)/T_{\rm eff} = (1/2)^{1/4} = 0.8409$. The compares well with the exact value $T(0)/T_{\rm eff} = 0.8114$.

Limb darkening

In the Eddington approximation we obtain the angle dependence of the emerging radiation by substituting $S(\tau) = 3H(\tau + 2/3)$ in eq. (4.53) and doing the integration. One finds

$$I(0,\mu) = S(\tau = \mu) = 3H\left(\mu + \frac{2}{3}\right) \tag{10.14}$$

For the relative run of total intensity over the stellar disk we then get

$$\frac{I(0,\mu)}{I(0,1)} = \frac{3}{5}\left(\mu + \frac{2}{3}\right), \tag{10.15}$$

where $I(0,1)$ is the intensity at the center of the stellar disk. The total intensity at the edge of the stellar disk ($\mu = 0$) is thus 40 percent of that at the center of the disk ($\mu = 1$). The effect is referred to as *limb darkening* and the result we have obtained is in first approximation in good agreement with the observed intensity change over the solar disk. It even agrees so well, that it spurred Karl Schwarzschild in 1914 to propose that the outer layers of the sun are in radiative equilibrium, and not in convective equilibrium as was expected on the basis of observations of the solar granulation (see § 9.5).

10.2 Constructing the grey atmosphere

Having obtained the temperature structure we still need to determine the pressure and density structure. In a grey atmosphere the equation of hydrostatic equilibrium (see eq. 9.32 and 9.35) reduces to

$$\frac{\mathrm{d}p_{\mathrm{G}}}{\mathrm{d}m} = \mathrm{g}_\star - \mathrm{g}_{\mathrm{R}}(m) = \mathrm{g}_\star - \frac{4\pi}{c}\,\chi' H = \mathrm{g}_\star - \frac{\sigma}{c}\,\chi'\,T_{\mathrm{eff}}^4 \tag{10.16}$$

For the relation between optical depth and column mass one gets

$$d\tau = (\chi/\rho)\,dm = \chi'\,dm \tag{10.17}$$

where χ' is the mean extinction coefficient in cm^2 gr^{-1}, which only depends on the abundances A_k, for all chemical elements k, and the state of the gas. For the latter we will assume LTE.

We define a reference point $z_\circ$ at the outer edge of the atmosphere, where the optical depth $\tau(z_\circ) \equiv \tau_\circ$ is small. We are free to choose the value of $\tau_\circ \ll 1$, for instance 10^{-6}. The temperature at this position follows from eq. (10.10) or (10.13) and will be almost identical to the limiting value $T(0)$. In order to determine the state of the gas we need to know, in addition to the temperature $T(\tau_\circ) \equiv T_\circ$, the density $\rho(\tau_\circ) \equiv \rho_\circ$. We now assume that beyond $z_\circ$, so even farther out, the state of the gas no longer changes. As in these outer layers, for all practical purposes, the temperature is constant, we obtain

$$\tau_\circ = \chi'_\circ \int_{z_\circ}^{\infty} \rho(z)\,dz = \chi'_\circ\,\rho_\circ\,H_\circ = \chi'_\circ\,m_\circ \tag{10.18}$$

where we have used eq. (9.30), substituting $\rho_\circ = \rho(z_\circ)$, and (9.34), and where the density scaleheight $H_\circ = \mathcal{R}T_\circ/\mu_\circ\,\mathrm{g}_{\mathrm{eff}}$, and $\mathrm{g}_{\mathrm{eff}} = \mathrm{g}_\star - \mathrm{g}_{\mathrm{R}}(m_\circ)$. The quantities $\chi'_\circ$ and $\mu_\circ$ are not known, after all, to know these requires knowledge of the state of the gas, for which we need $\rho_\circ$, which is the very quantity we aim to determine. This means we need to iterate for a moment at the position $z_\circ$ to get the correct $\rho_\circ$. Take as a starting solution an arbitrary mean extinction, say $\chi'_\circ$ = 0.5 cm^2 gr^{-1}. This yields by means of eq. (10.18) a value for the density $\rho_\circ$. Set the mean molecular weight $\mu_\circ$ equal to the mean atomic weight μ_{a} (see eq. 9.15). We can then compute the state of the gas from $N_\circ = \rho_\circ/(\mu_\circ m_{\mathrm{amu}})$ and $T_\circ$ following the method described in § 6.5. This results in new values for $\chi'_\circ$ and $\mu_\circ$, from which a new value of the density can be obtained, which, in turn, can be used to derive a new state of the gas, etcetera; until $\rho_\circ$ is converged. From the ideal gas law eq. (9.2) we find for the gas pressure at $m_\circ$

$$p_{\mathrm{G},\circ} = N_\circ k\,T_\circ. \tag{10.19}$$

We may now determine the structure of the entire atmosphere by integrating eq. (10.16) in the inward direction using an integration scheme for normal differential equations, such as for instance a 4$^{\mathrm{th}}$ order Runge-Kutta method. Use the value $\chi'_\circ$ to take a small step dm c.q. $d\tau$; this yields $p_{\mathrm{G}}(m_\circ + dm)$; determine using eq. (10.10) or (10.13) the value $T(m_\circ + dm)$; from

the ideal gas law we then find $N(m_\circ + dm)$, which gives us all we need to determine the state of the gas at $m_\circ + dm$. Take the next step, etcetera. Stop the integration if the total optical depth $\tau \sim 100$. This concludes the computation of the grey atmosphere.

However, there is one problem we have so far not discussed, which is how to determine a mean (i.e. grey) extinction coefficient for a gas of which the excitation and ionization state is known. This will be discussed in the next section.

10.3 Mean extinction coefficients

In a realistic situation, the extinction coefficient will be strongly frequency dependent. Especially so, when spectral lines are present. A grey extinction coefficient is therefore a quantity that needs to be defined. It implies that a grey atmosphere can at most give a – to some degree succesful or otherwise – approximation of reality. Still, it would be useful if we could preserve some of the results of the grey atmosphere problem, by making a careful choice of mean extinction coefficients. To start out, we restate the transfer equation and its first two moments in both the monochromatic and frequency-integrated grey case:

$$\begin{aligned} \mu\frac{\mathrm{d}I_\nu}{\mathrm{d}z} &= \chi_\nu(S_\nu - I_\nu) & \qquad \mu\frac{\mathrm{d}I}{\mathrm{d}z} &= \chi(J - I) \\ \frac{\mathrm{d}H_\nu}{\mathrm{d}z} &= \chi_\nu(S_\nu - J_\nu) & \qquad \frac{\mathrm{d}H}{\mathrm{d}z} &= 0 \\ \frac{\mathrm{d}K_\nu}{\mathrm{d}z} &= -\chi_\nu H_\nu & \qquad \frac{\mathrm{d}K}{\mathrm{d}z} &= -\chi H \end{aligned}$$

The monochromatic equations are on the left; the frequency-integrated grey equations are on the right. We will now discuss two means of defining a mean extinction coefficient, the flux-weighted mean, and the Rosseland mean.

Flux-weighted mean extinction

Say, we are interested in a mean extinction such that the 1^{st} order moment of the transfer equation reduces to the grey case. Integration of the monochromatic equation over frequency yields

$$-\frac{\mathrm{d}K}{\mathrm{d}z} = -\int_0^\infty \frac{\mathrm{d}K_\nu}{\mathrm{d}z}\,d\nu = \int_0^\infty \chi_\nu H_\nu\,d\nu = \chi_{\mathrm{F}} H \tag{10.20}$$

where

$$\chi_{\mathrm{F}} \equiv \frac{1}{H}\int_0^\infty \chi_\nu H_\nu\,d\nu \tag{10.21}$$

is the *flux-weighted mean extinction coefficient*. Note that the choice of $\chi_{\rm F}$ does not reduce the nongrey transfer problem completely to the grey case, for the transfer equation and its $0^{\rm th}$ order moment do not transform to their respective grey equivalents with this choice of mean extinction. Furthermore, there is the practical problem that H_ν is not known *a priori*, and therefore $\chi_{\rm F}$ cannot actually be calculated until after the transfer equation is solved. This latter difficulty can be overcome by an iteration between the construction of the grey model and the calculation of $\chi_{\rm F}$. Although the desired goal has not been fully attained, the fact that the flux-weighted mean preserves the K-integral is important, for it implies that the correct value is recovered for the radiation pressure $p_{\rm R}$, as well as for the radiation force $dp_{\rm R}/dz = {\rm g}_{\rm R}(z)$. This is of relevance for the determination of the density structure from the equation of hydrostatic equilibrium (see e.g. eq. 10.16).

Rosseland mean extinction

Say, we are interested in a mean extinction such that the correct value of the integrated flux H is conserved. That is, that we fulfill the constraint of radiative equilibrium. Using the $1^{\rm st}$ order moment of the transfer equation we obtain

$$-\int_0^\infty \frac{1}{\chi_\nu}\frac{{\rm d}K_\nu}{{\rm d}z}\,d\nu = \int_0^\infty H_\nu\,d\nu = H = -\frac{1}{\chi}\frac{{\rm d}K}{{\rm d}z} \tag{10.22}$$

where

$$\frac{1}{\chi} = \frac{\int_0^\infty (1/\chi_\nu)\,(dK_\nu/dz)\,d\nu}{\int_0^\infty (dK_\nu/dz)\,d\nu} \tag{10.23}$$

Again we face the practical difficulty that K_ν is not known *a priori*, and therefore χ defined according to the above definition can not be determined until the transfer equation is solved.

Still, it can be meaningful to use this definition of the mean extinction. However, we need to make an assumption. To do so, we use that at great optical depth (i.e. $\tau_\nu \gg 1$), where the radiation field is almost isotropic and the properties of the material medium are (locally) very close to thermodynamic equilibrium, $K_\nu \rightarrow 1/3 J_\nu$ (see § 3.6), and $J_\nu \rightarrow B_\nu$ (see § 6.2). In that case

$$\frac{{\rm d}K_\nu}{{\rm d}z} \simeq \frac{1}{3}\frac{{\rm d}B_\nu}{{\rm d}z} = \frac{1}{3}\frac{{\rm d}B_\nu}{{\rm d}T}\frac{{\rm d}T}{{\rm d}z} \tag{10.24}$$

which yields for the mean extinction

$$\frac{1}{\chi_{\rm R}} \equiv \frac{\int_0^\infty (1/\chi_\nu)\,(dB_\nu/dT)\,d\nu}{\int_0^\infty (dB_\nu/dT)\,d\nu} \tag{10.25}$$

This is the *Rosseland mean extinction coefficient*, which was already introduced in § 4.6 in the derivation of the diffusion approximation (see eq. 4.66). It is a harmonic mean, i.e. the largest contributions are from frequency regions where χ_ν is smallest and the transported flux is largest. The use of the Rosseland mean optical depth scale $\tau_{\rm R}$ will, at great optical depth, recover the correct shape of the asymptotic transfer equation (eq. 4.65), and, therefore, the

correct total flux (eq. 4.68). In regions of large optical depth the temperature structure of a stellar atmosphere in radiative equilibrium is thus given by

$$T^4(\tau_{\rm R}) = \frac{3}{4} T_{\rm eff}^4 \left(\tau_{\rm R} + q(\tau_{\rm R})\right) \tag{10.26}$$

For small optical depths the above approximation is no longer valid, as flux conservation near the stellar surface can no longer be guaranteed. Consequently, the temperature structure will deviate from eq. (10.26).

Exercise 10.1

Which fraction of the mass of the sun is in its atmosphere? Assume that the sun has an isothermal gray atmosphere for which $\chi'(z) = 0.8\ \mathrm{cm}^2\ \mathrm{gr}^{-1}$. We define the atmosphere to reach down to $\tau = 20$.

Exercise 10.2

Limb-darkening law Eq. (10.15) provides the specific intensity for angle μ relative to the value $I(\tau = 0, \mu = 1)$ at the center of the stellar disk.

a) Use the grey equivalent of Eq. (3.28) to compute the emergent grey flux as a function of $I(0, 1)$.

b) For which value of μ can the emergent grey flux be expressed as $\mathcal{F}(\tau = 0) = \pi I(0, \mu)$?

c) You had of course expected the result obtained in *b*? Why?

Exercise 10.3

Show that at large optical depth eq. (10.26) reduces to diffusion equation 4.68. Use table 10.1 to determine the behavior of $q(\tau_{\mathrm{R}})$ for $\tau_{\mathrm{R}} \gg 1$.

Exercise 10.4

Ponder on the results in this chapter and write down an expression for the Rosseland optical depth from the surface to the center of a star. Estimate the total radial Rosseland optical depth for the Sun, using information from the internet if needed.

11

LTE model atmospheres

In this chapter we discuss the construction of the LTE model atmosphere. The field of LTE atmospheres is dominated by the computer code ATLAS/SYNTHE and associated model grids of Robert L. (Bob) Kurucz [1]. We will therefore limit the discussion of LTE to this grid. Be ware, however, that a number of other independent codes exist; see Section 11.3 for a short overview.

11.1 Constructing the LTE atmosphere

We assume that the atmosphere consists of planar layers and that both hydrostatic and radiative equilibrium hold, i.e. we disregard convective energy transport which becomes important at temperatures below ~ 9000 K. The relevant transfer equation is eq. (4.45), which we repeat for clarity

$$\frac{\mathrm{d}^2(f_\nu J_\nu)}{\mathrm{d}\tau_\nu^2} = J_\nu - S_\nu = \frac{\kappa_\nu}{\kappa_\nu + \sigma_\nu}(J_\nu - B_\nu) \tag{11.1}$$

The second equality follows after substitution of the proto-type source function eq. (9.43). The extinction coefficient κ_ν is here the sum over all bound-bound, bound-free, and free-free absorption cross sections that may occur; σ_ν is the sum over all scattering processes. Recall that we derived the left part of this equation using the moments of the transfer equation. Note that we could also have obtained this result by integrating the 2^{nd} order differential equation for the symmetric average (eq. 5.17) over all solid angles. (In view of the definition of the symmetric average this integration would have been over the angles $0 \leq \mu \leq 1$). With this in mind, it is straightforward to derive the boundary conditions of eq. (11.1), i.e., we only need to integrate the boundary conditions of the symmetric average eq. (5.20) and (5.21) in a similar way. After having multiplied by μ, we obtain for the condition at the outer edge $\tau_\nu = 0$

$$\left.\frac{\mathrm{d}(f_\nu J_\nu)}{\mathrm{d}\tau_\nu}\right|_0 = g_\nu(0) J_\nu(0) \tag{11.2}$$

[1] For how to download the model atmosphere code ATLAS and spectral synthesis code SYNTHE and a python wrapper for ATLAS/SYNTHE, see the github package VidmaPy by Tomasz Różański

where the Eddington factor $g_\nu(0) = H_\nu(0)/J_\nu(0)$, which we, inspired by eq. (10.9) for the grey problem, set to the initial value $1/\sqrt{3}$ (see table 10.1). For the inner edge $\tau_\nu = \tau_{\max}$ we find

$$\left.\frac{\mathrm{d}(f_\nu J_\nu)}{\mathrm{d}\tau_\nu}\right|_{\tau_{\max}} = \frac{1}{3}\left(\frac{1}{\chi_\nu}\left|\frac{\mathrm{d}B_\nu}{\mathrm{d}z}\right|\right)_{\tau_{\max}} \tag{11.3}$$

Discretization of these equations is done using the finite difference method as explained in § 5. In eq. (11.1) we obviously choose the outermost right-hand-side term (i.e. not middle one) for discretization, as this allows us to account explicitely for the scattering term in the source function.

The radiation quantities for which we solve directly are J_ν and f_ν, therefore we choose to write the equation of hydrostatic equilibrium as

$$\frac{\mathrm{d}p_{\mathrm{G}}}{\mathrm{d}m} = \mathrm{g}_\star - \frac{4\pi}{c}\int_0^\infty \frac{\mathrm{d}(f_\nu J_\nu)}{\mathrm{d}m}\,d\nu \tag{11.4}$$

where we have used eq. (3.34), (3.36), (3.37) and (9.32). Discretization of this 1^{st} order differential equation yields

$$N_d kT_d - N_{d-1}kT_{d-1} + \frac{4\pi}{c}\sum_{n=0}^{N} w_n\left(f_{dn}J_{dn} - f_{d-1,n}J_{d-1,n}\right) = \mathrm{g}_\star\left(m_d - m_{d-1}\right) \tag{11.5}$$

where $\{m_d\}, d = 0, ..., D$ is the set of column masses, such that $d = 0$ describes the outer edge of the atmosphere, and $d = D$ the inner edge; w_n are the frequency integration weights of the set $\{\nu_n\}, n = 0, ..., N$ frequency points.

If we use the grey result (eq. 10.10) as a starting value of the temperature structure, to which we refer as $T_0(m)$, we may determine an initial value for the total particle density $N_0(m)$ analogously to the way we have done this in the grey problem (see § 10.2). The only difference is that we now need to assume a starting value for the mean intensity, for which we take $J_\nu(m) = B_\nu(T_0(m))$, and for the Eddington factor, for which we assume $f_\nu(m) = 1/3$. At each grid point we determine the state of the gas, based on the local values N_d and T_d. This then allows us to compute the extinctions κ_{dn} and σ_{dn}. Once we have derived the structure in this way, the radiation field J_{dn} can be derived from a solution of the transfer equation using the current values f_{dn}. These Eddington factors can then be updated by doing a formal solution, using first the new J_{dn} to update the total source function. This implies that we need to solve eq. (5.17) for the symmetric mean u_{dmn} for each combination of frequency n and angle of incidence m. The new Eddington factors follow from

$$f_{dn} = \frac{K_{dn}}{J_{dn}} = \frac{\sum_{m=0}^{M} w_m \mu_m^2 u_{dmn}}{\sum_{m=0}^{M} w_m u_{dmn}}, \tag{11.6}$$

where w_m are the angle integration weights of the set $\{\mu_m\}, m = 0, ..., M$. In a similar way we may determine new values for g_{0n}, required for the boundary condition (11.2).

The new mean intensities are used to test the equation of radiative equilibrium. In LTE this is eq. (9.44), which we here repeat

$$4\pi \int_0^\infty \kappa_\nu(z) \left[B_\nu(z) - J_\nu(z)\right] d\nu = 0 \tag{11.7}$$

and which in discrete form is given by

$$4\pi \sum_{n=0}^{N} w_n \kappa_{dn} (B_{dn} - J_{dn}) = 0 \tag{11.8}$$

where again w_n are the frequency integration weights of the set $\{\nu_n\}, n = 0, ..., N$ frequency points. From this equation we determine a new value for the temperature T_d at each grid point. How to do this is discussed in the next section. With this new run of temperatures we redo the above described process, and we keep iterating until T_d is converged. This gives the LTE model atmosphere.

Finally, the emerging spectrum of the LTE model is given by

$$\mathcal{F}_\nu(0) = 4\pi H_\nu(0) = 4\pi\, g_\nu(0) J_\nu(0) \tag{11.9}$$

The general source function for a gas

At the start of this section we have rather absentmindedly adopted the proto-type formalism (9.43) for the total source function. But, what is the correct expression for the total source function of a gas anyway?

From definition (4.32) we know that the total source function is nothing but the sum of all emission processes divided by the sum of all extinction processes. In the most general case we find for the total non-LTE extinction per centimeter

$$\begin{aligned} \chi_\nu \;=\; & \sum_k \sum_j \left\{ \sum_l \left\{ \sum_{u>l} \left[n_{ljk} - \frac{g_{ljk}}{g_{ujk}} n_{ujk} \right] \alpha_{lu,jk}(\nu) + \right. \right. \\ & \left. \left. \left[n_{ljk} - n^*_{ljk} e^{-h\nu/kT} \right] \alpha_{l,j+1,k}(\nu) \right\} + \alpha_{jk}(\nu, T) n_{\rm e} N_{jk} \left(1 - e^{-h\nu/kT} \right) \right\} + n_{\rm e} \sigma_{\rm T} \end{aligned} \tag{11.10}$$

where $\alpha_{lu,jk}$ and α_{ljk} are the extinction coefficients per particle, given by eq. (7.21) and (8.1), respectively; α_{jk} is defined as the leading part of eq. (8.17). The summation is over all elements k, ionization states j and excitation states l. Here u denotes a higher excitation state then l. The first term gives the line extinction, corrected for stimulated emission. The second term represents the bound-free extinction, again corrected for stimulated emission. Why for this last process the LTE population is used is immediately clear if one subtracts eq. (8.10) from (8.1). The third term gives the extinction by free-free transitions and the last term is that of electron scattering.

For the total non-LTE thermal emission coefficient (dimensions erg cm^{-3} sec^{-1} hz^{-1} sr^{-1}) it follows that

$$\eta_\nu = \frac{2h\nu^3}{c^2}\sum_k\sum_j\left\{\sum_l\left\{\sum_{u>l} n_{ujk}\frac{g_{ljk}}{g_{ujk}}\alpha_{lu,jk}(\nu)+ n^*_{ljk}\alpha_{l,j+1,k}(\nu)e^{-h\nu/kT}\right\} + n_{\rm e}N_{jk}\alpha_{jk}(\nu,T)e^{-h\nu/kT}\right\} \tag{11.11}$$

The three terms again describe bound-bound (i.e. line), bound-free, and free-free transitions. The emission by scattering is treated seperately in the proto-type source function (and is given by $\eta_\nu = n_{\rm e}\sigma_{\rm T}J_\nu$ for an isotropic source function) and is therefore not included in η_ν.

LTE source function for a gas

If LTE holds, then the extinction and thermal emission per centimeter is given by

$$\chi^*_\nu = \sum_k\sum_j\left\{\sum_l\left\{\sum_{u>l} n^*_{ljk}\alpha_{lu,jk}(\nu) + n^*_{ljk}\alpha_{l,j+1,k}(\nu)\right\} + n_{\rm e}N_{jk}\alpha_{jk}(\nu,T)\right\} \tag{11.12}$$

$$\times\left(1-e^{-h\nu/kT}\right) + n_{\rm e}\sigma_{\rm T} \tag{11.13}$$

$$\eta^*_\nu = \frac{2h\nu^3}{c^2}e^{-h\nu/kT}\times \tag{11.14}$$

$$\sum_k\sum_j\left\{\sum_l\left\{\sum_{u>l} n^*_{ljk}\alpha_{lu,jk}(\nu) + n^*_{ljk}\alpha_{l,j+1,k}(\nu)\right\} + n_{\rm e}N_{jk}\alpha_{jk}(\nu,T)\right\} \tag{11.15}$$

That indeed this is equivalent to the proto-type source function eq. (9.43) and that the thermal part of the source function fulfils the Kirchhoff-Planck relation (6.3) will be investigated in exercise 11.1.

11.2 Obtaining the temperature structure

The determination of the temperature structure is, in fact, the very heart of the problem of constructing LTE models. As discussed, we adopt the grey solution as the initial estimate of $T(m)$ in the LTE case. If we substitute this initial solution in eq. (11.7) we will normally find that the newly determined radiation field (the solution of eq. 11.1, 11.2, and 11.3) does not fulfil the constraint of radiative equilibrium. It is therefore necessary to adjust $T(m)$ iteratively in such a way that the radiation field ultimately satisfies the requirement of energy balance. There are basically two strategies we may use to make this happen: (A) temperature correction procedures, and (B) to intertwine the solution of radiative equilibrium and the transfer equation. We will discuss the basic principle of each family of methods.

(a) Lambda iteration

The solution of the mean intensity for frequency ν from the current value of the temperature $T(m)$ can formally be written as

$$\mathbf{T}_{\tau_\nu}\mathbf{J}_\nu = \mathbf{B}_\nu \tag{11.16}$$

where $\mathbf{T}_{\tau_\nu}$ is the $(D \times D)$ transfer matrix; $\mathbf{J}_\nu$ is the vector of mean intensities for all depth points d, and $\mathbf{B}_\nu$ is the vector of local Planck functions (given by the temperature $T(m)$). The transfer matrix can be constructed by discretization of eq. (11.1), and of its boundary conditions (11.2) and (11.3). If we adopt the central difference representation (such as described in section 5.2) we obtain for $\mathbf{T}_{\tau_\nu}$ a tri-diagonal matrix *à la* eq. (5.30). Inversion of this equation yields

$$\mathbf{J}_\nu = \mathbf{T}_{\tau_\nu}^{-1}\mathbf{B}_\nu = \mathbf{\Lambda}_{\tau_\nu}\mathbf{B}_\nu \tag{11.17}$$

where $\mathbf{\Lambda}_{\tau_\nu}$ is the *lambda matrix* or *lambda operator*. The iteration scheme described in section 11.1 is therefore often refered to as *lambda iteration* because, formally, the mean intensity is obtained by applying the $\mathbf{\Lambda}_{\tau_\nu}$ operator to the current $T(m)$, yielding J_ν, which in turn is used to improve $T(m)$.

The most obvious method of lambda iteration is to improve the temperature after each iteration step in the following way. We assume that the requirement of radiative equilibrium is fulfilled if we correct the current temperature $T(m)$ by $\Delta T(m)$, such that

$$4\pi \int_0^\infty \kappa_\nu(T)\left[B_\nu(T(m) + \Delta T(m)) - J_\nu\right] d\nu = 0 \tag{11.18}$$

If we expand the Planck function to within first order, i.e.

$$B_\nu(T + \Delta T) \simeq B_\nu(T) + \Delta T \left.\frac{\partial B_\nu}{\partial T}\right|_T \tag{11.19}$$

we find

$$\Delta T = \frac{\int_0^\infty \kappa_\nu(T)\left[J_\nu - B_\nu(T)\right] d\nu}{\int_0^\infty \kappa_\nu(T)(\partial B_\nu/\partial T)_T \, d\nu} \tag{11.20}$$

It must be emphasized that the value of J_ν in these equations denotes the value computed from the old $T(m)$ values. If one carries through this process and recomputes a new model with the new temperature distribution, one usually obtains some improvement in satisfying the requirement of radiative energy conservation. However, one often has to pass through hell to achieve convergence. The main reasons for this are: (1) It is only with difficulty that information on the non-local nature of the radiation field can penetrate regions of high optical depth, after all it is damped by the term $\exp(-\tau_\nu)$. If the initial values for $T(m)$ are far from the actual solution, it will typically take on the order of $\tau_\nu(D)$ iterations to propagate this information (and, to be more specific: this is information about the boundary conditions of the problem) through the medium. It takes so long for this information to seep through because the mean optical photon path by which the radiation field can make itself be known (per solution of the transfer equation, so per iteration) is only $< \tau_\nu >= 1$ (see eq. 4.27). Therefore,

if the solution of $J_\nu(m)$ is vastly different from the local value of the Planck function, it will take on the order of $\tau_\nu(D)$ iterations to correct for this. (2) The temperature correction $\Delta T(m)$ computed in the lambda iteration procedure may be complete nonsense because the effect of the correction does not (ddirectly) affect the values $J_\nu(m')$, i.e. the radiation field elsewhere in the medium. After all, the mean intensity is assumed constant when computing the temperature corrections.

The need for other methods to solve the LTE atmosphere problem (and other transfer problems) appears obvious. One very succesful alternative method was first introduced in astrophysics by C.J.Cannon (1973a, 1973b). This is the *method of postponed corrections*, better known as *operator splitting* or *approximate lambda iteration*.

(b) Approximate lambda iteration

At the basis of the approximate lambda iteration is the realization that some parts of the physical coupling between medium and radiation field are more important than others. The method aims to split the problem at hand in such a way that the problematic regions (those where the optical depth is large) are seperated from those where no problems occur (the optically thin regimes). The trick is to weave together the requirement of radiative equilibrium and the solution of the transfer equation in those regions where problems occur. In other parts of the medium we do not do anything special, we simply use the lambda iteration technique to come to a solution. Because the optical depth in the problematic regimes is large (per definition) it implies that the trick solution should act more or less *locally*. The splitting therefore focuses on defining the *local* and *non-local* medium, and the separation of the two. So, lets do that.

The formal solution of the mean intensity can be considered as a linear operation on the source function, i.e.

$$J_\nu = \Lambda_\nu S_\nu \tag{11.21}$$

Note that for this more general discussion we have chosen to have the lambda operator act on S_ν (see the first equality in eq. 11.1) and not B_ν (see the second equality in the same equation). This will be repaired once we return to the LTE atmosphere problem. We now formulate the idea of the splitting as follows

$$\Lambda_\nu = \Lambda^\star_\nu + (\Lambda_\nu - \Lambda^\star_\nu) \tag{11.22}$$

where $\Lambda^\star_\nu$ is an appropriately choosen *approximate lambda operator* which acts on the state of the medium and/or radiation field that still has to be determined. The operator $(\Lambda_\nu - \Lambda^\star_\nu)$ acts on the known, i.e. current state (in our LTE atmosphere problem this is the temperature $T(m)$). In short

$$J^{\rm new}_\nu = \Lambda^\star_\nu S^{\rm new}_\nu + (\Lambda_\nu - \Lambda^\star_\nu) S^{\rm old}_\nu \tag{11.23}$$

where "new" denotes that the source function is dependent on the state of the medium that still has to be determined, and "old" refers to the state of the medium as determined in the previous iteration. Note that *any* choice of $\Lambda^\star_\nu$ that leads to convergence will provide the

correct physical solution. After all, if $S_\nu^{\rm new} \rightarrow S_\nu^{\rm old}$ then $J_\nu \rightarrow \Lambda_\nu S_\nu$. To actually achieve convergence, as discussed above, $\Lambda_\nu^\star$ in regions of large optical depth should (locally) be a good approximation of the exact operator Λ_ν. Beware: not necessarily the best appromation, good is in principle good enough.

Based on the above discussion it may be clear that for this problem the ideal choice of a strictly local operator is

$$\Lambda_\nu^\star \equiv {\rm diag}\left[\mathbf{T}_{\tau_\nu}^{-1}\right], \tag{11.24}$$

for a diagonal matrix acts on the local source function. This definition has been proposed in 1986 by Olson, Auer, & Buchler and is often called the *OAB operator*. The problem with this definition seems, at first sight, that the explicit determination of this approximate lambda operator is cumbersome and computationally expensive, as it needs inversion of the transfer matrix $\mathbf{T}_{\tau_\nu}$, which requires order D^3 computations. However, Rybicki & Hummer (1991) have shown that for a tri-diagonal matrix system this type of problem can be solved in order D computations if one combines the forward recursive sweep and back substitution (see section 5.2) with a backward recursive sweep and back substitution.

We now return to the problem of LTE atmospheres (and switch back to using the second equality in eq. 11.1). With the new radiation field given by

$$J_\nu^{\rm new} = \Lambda_\nu^\star B_\nu^{\rm new} + (\Lambda_\nu - \Lambda_\nu^\star) B_\nu^{\rm old} = \Lambda_\nu^\star B_\nu^{\rm new} + \Delta J_\nu^{\rm old} \tag{11.25}$$

where $B_\nu^{\rm new} = B_\nu(T(m) + \Delta T(m))$ and $B_\nu^{\rm old} = B_\nu(T(m))$, one finds after substitution in eq. (11.7)

$$4\pi \int_0^\infty \kappa_\nu \left[(1-\Lambda_\nu^\star) B_\nu(T(m) + \Delta T(m)) - \Delta J_\nu^{\rm old}\right] d\nu = 0 \tag{11.26}$$

The term $\Delta J_\nu^{\rm old}$ describes the non-local contribution to the mean intensity; the local contribution to J_ν is, by means of this operator splitting, computed explicitly on the basis of the "correct" temperature $T(m) + \Delta T$. The constraint of radiative equilibrium is in this way (partly) woven together with the solution of the equation of transfer. The temperature corrections $\Delta T(m)$ again follow using a first order expansion of the Planck function (see eq. 11.19 and 11.20). We find

$$\Delta T = \frac{\int_0^\infty \kappa_\nu(T) \left[\Delta J_\nu^{\rm old} - (1-\Lambda_\nu^\star) B_\nu(T)\right] \, d\nu}{\int_0^\infty \kappa_\nu(T) (\partial B_\nu / \partial T)_T \, d\nu} \tag{11.27}$$

The convergence properties of this approximate lambda iteration method are excellent. The non-local radiation field, which in optically thick parts of the medium is only a small fraction of the total radiation field, is isolated and "drives" as it were the solution towards convergence. This is essential as it is this non-local radiation field that has to ensure that all points in the medium become aware of the boundary conditions. Locally the radiation field is determined in compliance with radiative energy conservation. In this way of dealing with the problem the corrections $\Delta T(m)$ are reliable (recall that this need not be the case in standard lambda iteration).

Other definitions of $\Lambda^\star_\nu$ have been proposed in order to solve transfer problems. Examples are the *core saturation operator* and the *Scharmer operator*, both introduced by Scharmer in 1981.

To actually compute detailed LTE model atmospheres we obviously need to know how to describe the spectral lines in detail. This we will discuss in the next chapter. In anticipation of this discussion we already discuss aspects of the end result in the next section.

11.3 LTE modeling codes: ATLAS, MARCS, PHOENIX, TLUSTY

The most renowned LTE codes are the planar ATLAS/SYNTHE model by Bob Kurucz and the spherical MARCS model by the Scandinavian group led by Bengt Gustafsson. Note that most of the NLTE codes can be used to calculate LTE models as well, e.g. PHOENIX and TLUSTY.

An extensive grid of LTE models is that of Kurucz and we will discuss it in some detail below. A grid of 52 000 LTE MARCS models is computed by Gustafsson et al. (2008, A&A 486, 951). These span the effective temperature range 2500 K to 8000 K, surface gravity range $-1 \leq \log g \leq 5$, and metallicity range $-5 \leq [\mathrm{Metal/H}] \leq +1$. A main difference with the Kurucz models is the much more comprehensive treatment of molecular opacities. PHOENIX, developed by the group of P. (Peter) Hauschildt, can be used to calculate atmospheres and emerging spectra of stars all across the cooler part of the HR-diagram, including cool giants, brown dwarfs and extrasolar giant planets. PHOENIX also has the option to compute (partial) non-LTE models. An extensive library of PHOENIX stellar atmospheres and synthetic spectra is also available. The parameter space of this grid covers 2300 K to 12000 K, surface gravity $0.0 \leq \log g \leq 6.0$, and metallicity $-4.0 \leq [\mathrm{Metal/H}] + 1.0$. TLUSTY/SYNSPEC, by Ivan Hubeny and Thierry Lanz, is mostly used for hot stars, where NLTE is important. See also Hubeny & Lanz (1995) and Hubeny et al. (2021) for details.

For an overview of hot star NLTE models that account for stellar winds, see Section 15.6 and Table 15.1.

Kurucz models

In the Kurucz models it is assumed that LTE and hydrostatic equilibrium are valid, and that the energy is transported by either radiation or convection. In the final set of models the contributions of over 58,000,000 million spectral lines have been taken into account.

A very large standard grid op Kurucz models is available, in which for 19 different sets of abundances, 76 effective temperatures and 11 gravities, the atmospheric structure and emerging spectrum have been computed (so a total of 15 884 models). The specification of the abundance pattern is always such that the total mass fraction of all elements more heavy than helium, the so-called *metal abundance* Z, are scaled to the solar abundances $Z_\odot = 0.018$. In

this way the $\log Z/Z_\odot$ value is varied between +1.0 and -5.0. The values for the effective temperature range between $3\,500 \leq T_{\rm eff} \leq 50\,000$ K, and that of the logarithm of the gravity between $0.0 \leq \log {\rm g} \leq 5.0$.

An extensive library of synthetic spectra based on Kurucz's code that covers the 2500-10500 Å wavelength range at resolving powers (see Eq. 13.3) 20,000, 11,500, 8,500 and 2,000 has been presented by Munari et al. (2005). Their spectra are available as absolute fluxes as well as continuum normalized fluxes.

The purpose of these models is to derive the important quantities $T_{\rm eff}$, $\log {\rm g}$, and the chemical abundance pattern by comparing them to observed energy distributions. Also, these models provide colors, bolometric corrections, ionizing fluxes – relevant for nebular studies – and limb darkening profiles – relevant for studies of the light curves of eclipsing binaries and exo-planet transits, of line profiles of rotating stars and proper interpretation of measurements of stellar radii using imaging and interferometric techniques. After correcting for interstellar extinction (see § 19.2), scaling of an appropriate model with an absolute measurement of the flux (for instance the V magnitude or a part of the spectrum for which an absolute calibration is available) then yields a value for the angular diameter (see eq. 3.31). If we know the distance to the object, for instance from a parallax measurement, we then have a value for the stellar radius. Using eq. (3.25) this yields a value for the luminosity, and, with eq. (9.28), the current mass.

A comparison with evolutionairy tracks plotted in the Hertzsprung-Russell diagram of $L_\star$ versus $T_{\rm eff}$, then reveals the age, initial mass, and evolutionary stage of the star. The web interface BONNSAI offers an easy tool to do just this and more (see also Schneider et al. 2014).

We briefly discuss some aspects of the Kurucz models.

Atmospheric structure

Table B.9 gives for a number of effective temperatures $T_{\rm eff}$ and gravities $\log {\rm g}$ a sampling of the most important structural parameters as a function of the continuum optical depth at $\lambda 5000$ Å. A gravity $\log {\rm g} = 4$ is typical for main sequence stars (luminosity class V); $\log {\rm g} = 1$ is characteristic for the rarefied atmospheres of giants (luminosity class III). Note that e.g. the density ρ at $\tau \sim 1$ in the giant star is two orders of magnitude less than in the dwarf star. Deep down ($\tau \sim 10$) in the atmospheres of cool stars convection is the most important mode of energy transport; in higher layers – where the mean free path of the photons start to become large – energy is transported by radiation. In stars with $T_{\rm eff} \geq 10\,000$ convection does not occur anywhere in the atmosphere. In hotter stars hydrogen gets ionized, such that $n_{\rm e} \sim N_N$, where N_N is the H dominated number density of nuclei. In hot stars radiation pressure becomes so important that $p_{\rm R} > p_{\rm G}$.

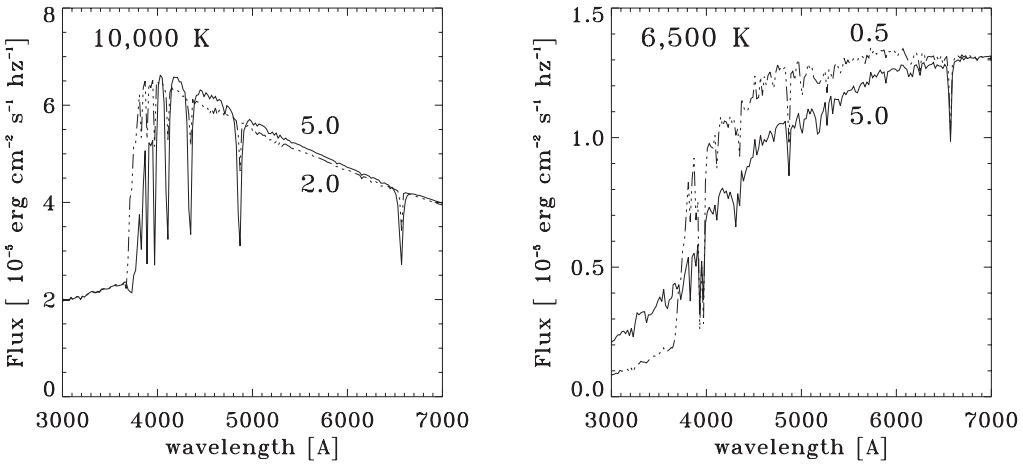


Figure 11.1: *Kurucz models with a solar abundance pattern and a micro-turbulent velocity of 2* $\mathrm{km\,sec^{-1}}$ *for 10 000 K (left panel) and 6 500 K (right panel) and different values of the gravity.* $\log g$*= 5.0 (that could be a dwarf star); 2.0 (a bright giant or supergiant), en 0.5 (a supergiant).*

Energy distribution

Figure 11.2 shows examples of energy distributions computed by Kurucz. The wavelength resolution of these spectral energy distributions (SEDs) is 10 Å, i.e. too poor to properly resolve individual spectral lines. Still, some of the strongest lines can be identified, notably the Balmer series lines of hydrogen.

$U-B$ vs. $B-V$ diagram

The Kurucz color-color diagram of $U-B$ vs. $B-V$ is given in figure 11.3. The top panel gives this diagram on the same scale as that of figure 2.11 and 6.6; the bottom panel shows an enlargement of the regime where these indices are the most sensitive to gravity. Labels indicate the effective temperature and logarithm of the gravity of the grid of models. Notice that a star that has $T_{\rm eff}$ = 6 500 K is 'redder' in its $B-V$ color if the gravity is larger, but is 'bluer' in its $U-B$ color. This is a reflection of the behavior of the optical spectrum as a function of g. For the two extreme values of the gravity these spectra are plotted in the right panel of Fig. 11.1. Remember that the U-filter is positioned mainly in the Balmer continuum (i.e. at the short-wavelength side of 3646 Å; see Fig. 6.4) and the B- and the V-filter are not. The star with the high gravity has a relatively high flux in U and low flux in B; therefore a relatively 'blue' $U-B$ color. Comparison of the flux in the B- and V-filter shows that the high gravity model has (relatively speaking) less flux in the B-filter than in the V-filter; therefore a relatively 'red' $B-V$ color.

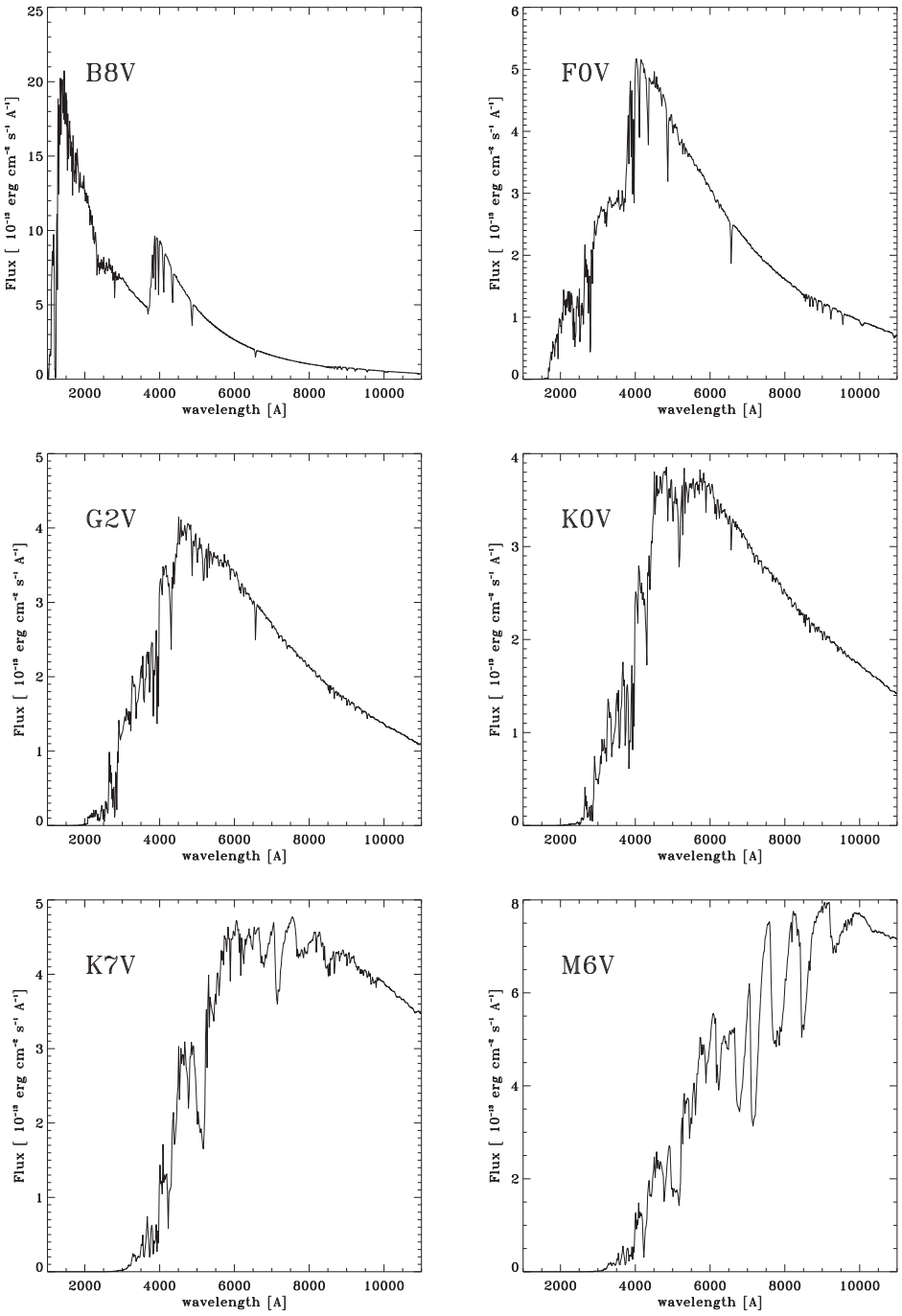


Figure 11.2: *Examples of the energy distributions of Kurucz models for mains sequence star with solar abundances.*

Kurucz colors and bolometric correction BC_V

An extensive set of synthetic colors can be found at one of the http addresses listed in the beginning of this subsection, in the folder *Grids of model atmospheres*. In 2008 opacities and abundances have been updated. Color information regarding $U-B$, $B-V$, $V-R$, $V-I$, $V-J$, $V-K$, $V-$L is supplied for various chemical compositions and a few microturbulent velocities. For instance, for a solar composition and $\xi_{\rm micro} = 2\,{\rm km\,sec^{-1}}$ the grid `P00ODFNEW` and tables `ubvp00k02odfnew.dat` and `rijklp00k2odfnew.dat` supply the relevant information. The former table also provides bolometric corrections BC_V in the V-band. Recall that

$$M_{\rm bol} = BC_f + M_f, \tag{11.28}$$

where M_f is the absolute magnitude of photometric filter f. The measured apparent magnitude m_f and M_f are related by Eq. 2.6.

The Kurucz colors make use of measured filter transmissions in the Johnson system. For the rest, they are fully synthetic in nature, i.e. they rely on the Kurucz models. Obviously, any model is susceptible to uncertainties. For the Kurucz models these include an incomplete line list, assumptions in the treatment of convection, turbulence and line broadening and the assumption of LTE. For this reason, others have determined colors using 'more' empirical approaches. For instance, Worthey & Lee (2011) use a catalog of *observed* photometric data of nearby stars that have been analyzed using (mainly) Kurucz models to yield $T_{\rm eff}$, $\log g$, and [Fe/H]. These tables, together with a fortran program that provides Johnson $U-B$, $B-V$, $V-R$, $V-I$, $J-K$, $H-K$, $V-K$ and the bolometric correction, are available online.

Useful relations between $B-V$ color and effective temperature and/or bolometric correction have been derived by several authors. For instance, Flower (1996) provides polynomial fits (see table 11.1) as well as an extensive table.

Limb darkening

The amount of information in the Kurucz-models on limb darkening, i.e. the run of specific intensity from the center of the stellar disk to the edge, is immense. This information is available for each viewing angle μ, for each frequency point. In practice this information is usually described with a limb-darkening law (that describes the angle dependence) for a number of photometric bands. The first to propose a limb-darkening law was Milne (1921)

$$\frac{I_m(\mu)}{I_m(1)} = 1 - u_m(1-\mu) \tag{11.29}$$

where $I_m(\mu)$ is the specific intensity in the photometric filter m (see § 2.6)

$$I_m(\mu) = \frac{\int_0^\infty I_\nu(\mu)\mathcal{S}_m(\nu)\,d\nu}{\int_0^\infty \mathcal{S}_m(\nu)\,d\nu}. \tag{11.30}$$

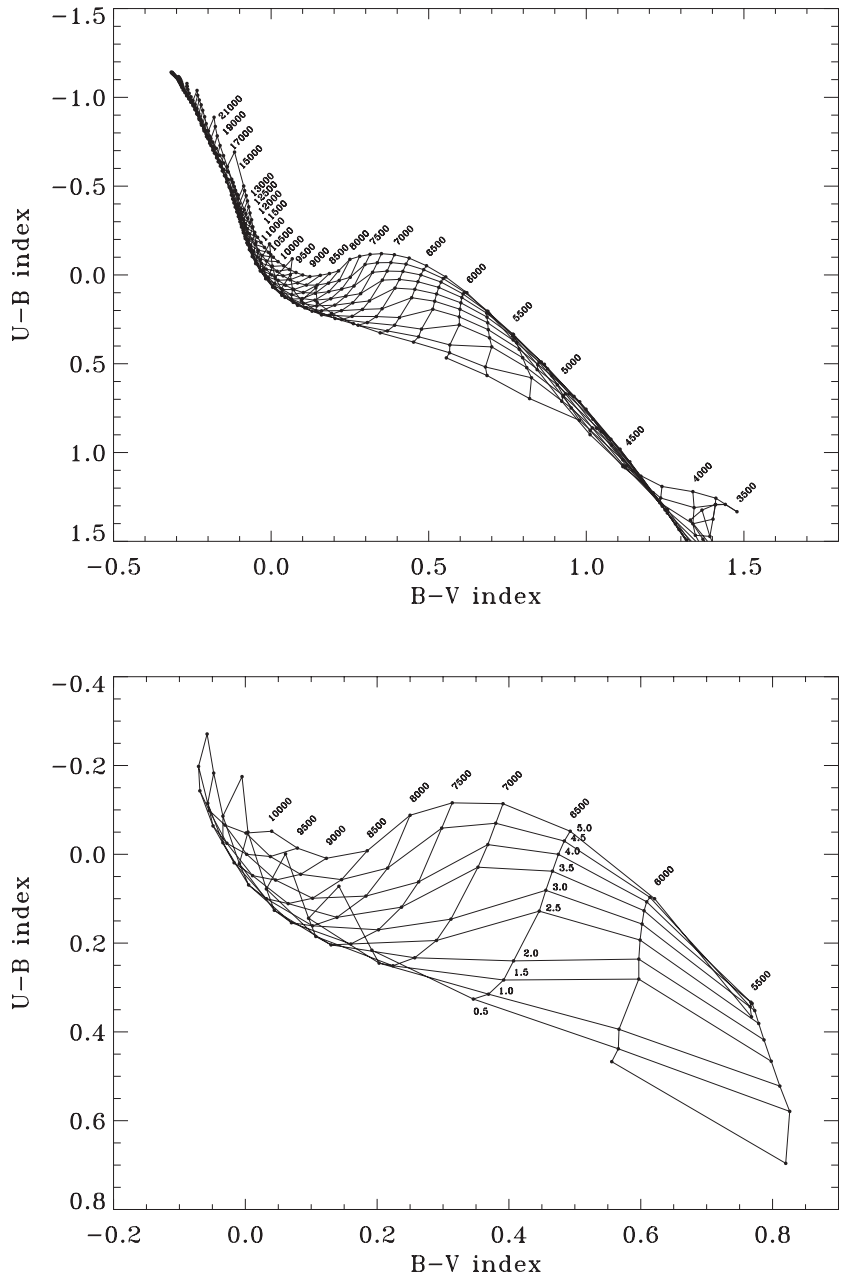


Figure 11.3: *Relation between the color indices $U-B$ and $B-V$ for Kurucz models with solar abundances and 2* $\mathrm{km\,sec^{-1}}$ *micro turbulent velocity. Top: global view. Bottom: detail. Labels provide* $T_{\rm eff}$ *and* $\log g$.

Table 11.1: *Polynomial relations expressing the $B-V$ color and BC_V as a function of effective temperature (Flower 1996). The $B-V$ relation for supergiants is slightly different from that of other luminosity classes.*

$B-V$ color $B-V = a + b\log T_{\rm eff} + c(\log T_{\rm eff})^2 + ..$			Bolometric Corrections $BC_V = a + b\log T_{\rm eff} + c(\log T_{\rm eff})^2 + ..$		
Coefficient	V, IV, III	I	$\log T_{\rm eff} > 3.90$	$3.90 < \log T_{\rm eff} < 3.70$	$\log T_{\rm eff} < 3.70$
a	3.979145	4.0125597	-0.188115	-0,370510	-0.190537
b	-0.654499	-1.055043	-0.137146	0,385673	0.155145
c	1.740690	2.133395	-0.636234	-0,150651	-0.421279
d	-4.608815	-2.459770	0.147413	0,261725	0.381476
e	6.792600	1.349424	-0.179587	-0,170624	–
f	-5.396910	-0.283943	0.788732	–	–
g	2.192970	–	–	–	–
h	-0.359496	–	–	–	–

$I_m(1)$ is the specific intensity at the center of the stellar disk. Milne thus proposed limb-darkening to be linear phenomenon. For a number of Kurucz-models the limb-darkening coefficients u_m are give in table B.8. This linear description turned out to be a fairly good approximation for solar-type stars. More accurate laws are the quadratic law: $I_m(\mu)/I_m(1) = 1 - a_m(1-\mu) - b_m(1-\mu)^2$; de square root law: $1 - c_m(1-\mu) - d_m(1-\sqrt{\mu})$; and the logarithmic law: $1 - e_m(1-\mu) - f_m\mu\ln\mu$. A very accurate description is given by Claret (2000, A&A 363, 1081)[2] for metallicities in the range $-5.0 \leq \log Z/Z_\odot \leq 1.0$

$$\begin{aligned} \frac{I_m(\mu)}{I_m(1)} &= 1 - a_{1,m}(1-\mu^{1/2}) - a_{2,m}(1-\mu) - a_{3,m}(1-\mu^{3/2}) - a_{4,m}(1-\mu^2) \\ &= 1 - \Sigma_{k=1}^{4} a_{k,m}(1-\mu^{k/2}) \end{aligned} \tag{11.31}$$

Limb darkening is extremely important in the study of the light curves of eclipsing binaries and transiting extra-solar planets, the determination of stellar diameters, and the analysis of line profiles from rotating stars.

[2]For the entire Kurucz grid all coefficients can be found at
`http://vizier.cfa.harvard.edu/viz-bin/VizieR?-source=J/A+A/363/1081`.

Exercise 11.1

a) Show that, using $\kappa_\nu^* = \chi_\nu^* - n_\mathrm{e}\sigma_\mathrm{T}$ one recovers $\eta_\nu^* = \kappa_\nu^* B_\nu(T)$, as expected from the Kirchhoff-Planck relation (6.3).

b) Show that in LTE the proto-type source function is given by eq. (9.43).

Exercise 11.2

In this computer exercise we make a model for the effect on the lightcurve of a star due to the transit of an exo-planet. We model the transit as an eclipse of a spherical star by an opaque, dark sphere. In what follows, a is the center-to-center distance between the star and the planet, R_p is is the radius of the planet; $R_\star$ is the stellar radius; $z = a/R_\star$ is the normalized separation of the centers, and $p = R_\mathrm{p}/R_\star$ is the size ratio of the two objects. The flux relative to the unobscured flux ${\mathcal{F}_\nu}^\star$ is (Mandel & Agol 2002)

$$\frac{\mathcal{F}_\nu(p.z)}{{\mathcal{F}_\nu}^\star} = 1 - \Psi(p,z), \tag{11.32}$$

where

$$\Psi(p,z) = \begin{cases} 0, & 1+p \le z, \\ \frac{1}{\pi}\left[p^2 e_0 + e_1 - \frac{1}{2}\sqrt{4z^2 - (1+z^2-p^2)^2}\right], & |1-p| < z \le 1+p, \\ p^2, & z \le 1-p, \\ 1, & z \le p-1, \end{cases} \tag{11.33}$$

and $e_0 = \cos^{-1}[(p^2+z^2-1)/2pz]$ en $e_1 = \cos^{-1}[(1-p^2+z^2)/2z]$.

a) Which four cases are described in eq. (11.33)

b) Give a relation that describes the effect of inclination. Use your computer program to graphically show the effect of inclination and p.

Now add the effect of limb darkening to your model. Assume that the specific intensity of the part of the stellar surface occulted by the planet disk is constant and given by the specific intensity at the location of the center of the planetary disk. This approximation is typically better than 2% of $1 - \Psi(p,0)$ for $p < 0.1$. Assume that during ingress and egress the specific intensity of the stellar surface blocked by the planetary disk is constant as well and given by the value halfway between the edge of the star and farthest point of the planet on the stellar disk.

c) Assume for the star $T_\mathrm{eff} = 5\,500$ and $\log g = 4.5$. Adopt a linear limb-darkening prescription (see table B.8) for the relevant limb-darkening coefficient. Show graphically the effect of inclination and p.

12

Planetary atmospheres

12.1 Introduction

Orbiting planets may have two sources of energy heating their atmospheres, a source of radiative illumination by the host star and a source of interior energy. The amount of radiation from the star reaching the planet surface is critical. It heats the planet, drives currents in the planet atmosphere and ultimately governs the global energy balance. The internal source itself consists of two components. First, a component due to the gradual loss of residual gravitational potential energy from the planet's formation. This source may be particularly important when the planet is highly inflated, young, or far from its host star. For instance, Jupiter, being quite far from the Sun, has an internal luminosity over twice as high as its luminosity from reradiated absorbed stellar energy. Second, a component due to radioactive decay of long-living isotopes of predominantly uranium, thorium, and potassium. Of the two internal sources, this one is the most important for Earth. Still, for many planets, including Earth and the hot exoplanets in short-period orbits, illumination by the host star overwhelms the interior flux. For a conceptual overview of the exoplanet atmosphere problem see, e.g., Heng & Marley (2017).

Before we embark on a discussion of the thermal and hydrostatic structure of exoplanet atmospheres, let us first introduce and discuss some elementary concepts of planets and planetary atmospheres.

Energy balance

As in stellar atmospheres, no energy is created or destroyed in planetary atmospheres. All the energy in the planet atmosphere is either supplied by the host star in the form of absorbed incident radiation or by the planetary interior. The energy balance can thus be described as

$$L_{\rm out}(t) = (1 - A_{\rm B})\, L_{\rm inc}(t) + L_{\rm int}(t), \tag{12.1}$$

where $L_{\rm out}$ is the energy per second leaving the planet, $L_{\rm inc}$ the stellar energy per second incident on the planet, and $L_{\rm int}$ the energy per second transferred to the atmosphere from the

planetary interior. The fraction of incident stellar energy scattered back into space is the *Bond Albedo* $A_{\rm B}$ (see below). Therefore, $(1-A_{\rm B})L_{\rm inc}$ is the part of the total incident stellar energy per second absorbed to heat the atmosphere or surface and $A_{\rm B}L_{\rm inc}$ the part reflected back into space.

For a planet of radius $R_{\rm p}$, the total energy incident per second on its surface is

$$L_{\rm inc} = \pi R_{\rm p}^2\, \mathcal{F}_\star(a) = \pi R_{\rm p}^2 \left(\frac{R_\star}{a}\right)^2 \mathcal{F}_\star(R_\star) = \pi R_{\rm p}^2 \left(\frac{R_\star}{a}\right)^2 \sigma\, T_{\rm eff,\star}^4 \tag{12.2}$$

where a is the orbital separation and $\mathcal{F}_\star(a)$ is the total stellar flux at that distance. If planets and their host stars are the focus of attention, it is custom to add a p subscript to planet quantities and a $\star$ subscript to stellar quantities.

The total irradiance at Earth (see also Sect. 3.4) is given by

$$\mathcal{R} = \frac{L_{\rm inc}}{\pi R_{\rm p}^2} = \left(\frac{R_\star}{a}\right)^2 \sigma\, T_{\rm eff,\star}^4 = 1.362 \times 10^6\ {\rm erg\,cm^{-2}\,sec^{-1}} = 1362\ {\rm W\,m^{-2}} \tag{12.3}$$

where we have used numbers from Table 16.1. By definition, the incoming radiation is impinging the dayside of our planet. The average incoming solar energy per unit area of the Earth's surface is $L_{\rm inc}/4\pi R_{\rm p}^2 = \mathcal{R}/4$.

Equilibrium temperature

The equilibrium temperature of a planet is the effective temperature of that planet assuming it has no internal luminosity. It is derived from equating the energy per second emitted by the planet with the energy per second absorbed by the planet

$$\frac{1}{f}\, 4\pi R_{\rm p}^2\, \sigma\, T_{\rm eq}^4 = (1-A_{\rm B})\, \pi R_{\rm p}^2 \left(\frac{R_\star}{a}\right)^2 \sigma\, T_{\rm eff,\star}^4 \tag{12.4}$$

Consequently

$$T_{\rm eq} = \left[f\,(1-A_{\rm B})\right]^{1/4} \left(\frac{R_\star}{2a}\right)^{1/2} T_{\rm eff,\star} = \left[f'\,(1-A_{\rm B})\right]^{1/4} \left(\frac{R_\star}{a}\right)^{1/2} T_\star \tag{12.5}$$

where $f = 4f'$ is a correction factor to the term 4π and describes the effectiveness of atmospheric circulation, i.e. the degree to which energy absorbed is transferred from the planet's dayside to nightside. In case of full heat circulation around the planet, $f = 1$. If the dayside alone reradiates the incident energy (and the nightside remains cold), its higher resulting equilibrium temperature is given by $f = 2$. Further adjustments of f are used to account for the fact that the angle of incidence of the stellar flux decreases from the substellar point to the terminator. For instance, if the atmosphere instantly reradiates the absorbed radiation (with no advection), $f = 8/3$. For simplicity of notation, we have introduced $T_\star = T_{\rm eff,\star}$.

Using $f = 1$ and the Bond Albedos given in Table 12.1, we find $T_{\rm eq} = 254\,{\rm K}$ for Earth and $T_{\rm eq} = 229\,{\rm K}$ for Venus.

Planetary albedos

The planetary albedo is a measure of the reflectivity of the planet's surface and/or atmosphere. One may define the albedo in several ways; here we focus on the conceptually simplest one – the Bond albedo. $A_{\rm B}$ is the fraction of incident luminosity scattered back into space by the planet. Therefore, $0 \leq A_{\rm B} \leq 1$. The Bond albedo includes back-scattered radiation in all directions and radiation at all frequencies. The *spherical albedo* $A_{\rm S}(\nu)$ is the monochromatic Bond albedo, hence

$$A_{\rm B} = \int_0^\infty A_{\rm S}\, d\nu \tag{12.6}$$

Be ware that both the spherical albedo and the Bond albedo depend on the spectrum of the host star, i.e. they are not intrinsic properties of the planet. They are essentially weighted by the incident radiation (see e.g. Del Genio et al. 2019. Imagine an atmosphere in which Rayleigh scattering is important. As, in that case, proportionally more red photons are absorbed than blue photons (which scatter 16 times as easily; see Sect. 8.4), an identical planet will have a lower $A_{\rm B}$ under the light of a red star than a blue one. Also take note that the Bond albedo may change during the evolution of the host star and exoplanet.

Object	Bond albedo
Mercury	0.088
Venus	0.76
Earth	0.306
Moon	0.11
Mars	0.25
Jupiter	0.503
Saturn	0.342
Enceladus	0.8
Uranus	0.300
Neptune	0.290
Pluto	0.4
Eris	0.96

Figure 12.1: *Bond albedos. From: Wikipedia lemma Bond albedo.*

A planet with a high $A_{\rm B}$ scatters efficiently, hence is 'bright' in the optical. Venus has a high $A_{\rm B}$ due to its complete coverage of H_2SO_4 clouds. The Saturnian satellite Enceladus too has a high Bond albedo due to water geysers constantly producing fresh ice layers, making its surface very reflective. Mercury and the moon, in contrast, have no atmosphere, clouds, or ice. This makes them 'dark', scattering only 0.1 of the incident radiation. Though the Bond albedo of Jupiter is appreciable, this appears not the case for hot extrasolar giant planets. For HD 209458b, which has $T_{\rm eff,p} = 1442$ K, it is found that $A_{\rm B} < 0.12$ (Rowe et al. 2008). This low albedo rules out the presence of bright reflective clouds in this exoplanet's atmosphere. In the absence of clouds, all hot Jupiter models predict extremely low visible-wavelength (geometric) albedos, due to strong, broad absorption lines of neutral atomic Na and K. These atoms are known to suppress the emitted visible-wavelength flux of brown dwarfs at similar temperatures (Liebert 2001). Hot exoplanets are exposed to a much more intense UV radiation field from the host star than is Jupiter. This strong incident UV flux can produce a rich mixture of compounds from molecules producing a haze that can absorb incident UV photons and will darken the appearance of the planet in the blue part of the spectrum (Marley et al. 1999).

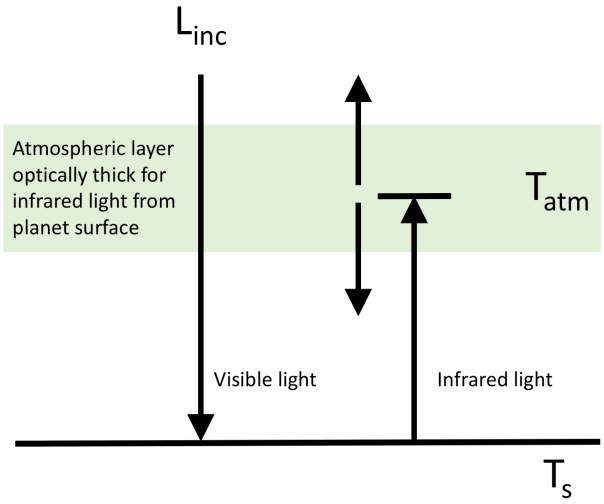


Figure 12.2: *A simple two-layer model to illustrate the difference between the atmosphere and surface temperature and the concept of the greenhouse effect.*

A simplified atmosphere – a toy model for the greenhouse effect

We consider a very simple model for a planetary atmosphere in which the atmosphere is represented by a single layer of temperature $T_{\rm atm}$ and that has a surface at temperature $T_{\rm s}$ (see Fig. 12.2). The absorbed stellar radiation is redistributed uniformly over the surface, i.e. $f = 1$. The atmosphere (1) is completely transparant to the incoming stellar radiation (so $A_{\rm B} = 0$). For solar-type stars the bulk of the incoming light is at optical wavelengths, so it is optical light that reaches the ground, is absorbed, heats the surface, and is reemitted at longer wavelengths. The atmosphere (2) is completely opaque to infrared radiation, i.e. it absorbs all the IR photons radiating from the ground. This is characteristic for a planetary atmospheres that is composed of mostly molecules as these absorb efficiently at IR wavelengths.

At its top, the opaque atmosphere radiates a flux $\sigma T_{\rm atm}^4$ out to space. At its bottom, given the symmetry of a homogeneous finite slab (see Sect. 4.6), the same amount is radiated toward the surface. From energy balance, the radiation emerging from the top of the atmosphere must be equivalent to the absorbed stellar radiation. Hence,

$$\sigma T_{\rm atm}^4 = (1 - A_{\rm B}) \left(\frac{R_\star}{2a} \right)^2 \sigma \, T_\star^4 = \sigma \, T_{\rm eq}^4 \tag{12.7}$$

At the surface, energy balance too means that all of the absorbed incoming radiation is again reradiated. There are two contributions to the absorbed radiation at the surface: the incoming stellar radiation and the downward radiation from the atmosphere layer. This yields

$$\sigma T_{\rm s}^4 = \sigma \, T_{\rm eq}^4 + \sigma T_{\rm atm}^4 = 2\sigma \, T_{\rm eq}^4 \tag{12.8}$$

This gives

$$T_{\rm s} = 2^{1/4}\, T_{\rm eq} \simeq 1.19\, T_{\rm eq} \tag{12.9}$$

Let us apply our model to Earth and Venus, for which the actual mean surface temperatures are 288 K and 730 K. Inserting the equilibrium temperatures computed above, we find $T_{\rm s} = 302\,{\rm K}$ for Earth and $T_{\rm s} = 273\,{\rm K}$ for Venus.

12.2 Vertical thermal structure of planetary atmospheres

If we assume the direction of the host star to be $(\theta_\circ, \phi_\circ)$ away from the surface normal, and the incident radiation to be plane parallel at the planet surface, then

$$I^\star_\nu(0,\mu) = I^\star_\nu\, \delta(\mu + \mu_\circ)\, \delta(\phi - \phi_\circ) \qquad \text{for } 0 \leq \mu_\circ \leq 1 \tag{12.10}$$

The $\star$-symbol denotes stellar quantities. Given our positive definition of $\mu_\circ$, it being the direction in which the star is seen from the atmosphere, the direction from which the light is coming is $-\mu_\circ$. Notice that because the host star illuminates the surface from a specific direction, axial symmetry (see Sect. 3.1) is lost. Below we discuss how one might deal with this added complexity. We may also write down the relationship between the incident stellar flux and the incident stellar specific intensity. Writing $\mathcal{F}^-_\nu = \mathcal{F}_\nu{}^\star$ in Eq. 3.21 we find

$$\mathcal{F}_\nu{}^\star(0) = \int_0^{2\pi} \int_0^{-1} \mu\, I^\star_\nu\, \delta(\mu + \mu_\circ)\, \delta(\phi - \phi_\circ)\, d\mu\, d\phi = \mu_\circ\, I^\star_\nu, \tag{12.11}$$

where $\mathcal{F}_\nu{}^\star(0)$ is positive. Notice that integration proceeds from a higher value to a lower value, which is why an extra negative sign enters. Integration over all frequencies yields

$$\mathcal{F}^\star = \mu_\circ\, I^\star, \tag{12.12}$$

where $I^\star$ is the total specific intensity (see Eq. 3.3) and $\mathcal{F}^\star$ the total incident stellar flux (see Eq. 3.23). In section (3.4), we have gone through some length to prove the $1/r^2$ dependence of the flux. The above equation implies a constant flux. How can that be? The reasons is that in the current case the incoming radiation is planar; see question (12.1).

12.2.1 Temperature structure - the grey atmosphere revisited

To obtain an expression for the vertical thermal structure of a planetary atmosphere we again resort to the grey, planar, LTE atmosphere where energy is transported only by radiation (see Chapter 10).

In radiative equilibrium the total flux passing through the atmosphere is constant. The radiative equilibrium temperature profile is the temperature profile that satisfies this constraint. The total flux that exits at the top of the atmosphere is

$$\mathcal{F} = \mathcal{F}_{\rm int} + \mu_\circ\, I^\star = \sigma\, T^4_{\rm int} + \mu_\circ\, I^\star, \tag{12.13}$$

where the first term is the total flux coming from the planetary interior that passes through the planet atmosphere. $T_{\rm int}$ is the interior temperature. To be clear about this: the *net* flux through the atmosphere is still only the internally generated flux $\mathcal{F}_{\rm int}$, since no energy from the irradiation is permanently absorbed by the planet – it is all re-radiated.

12.2.2 Atmosphere heated from below

Let us first consider an atmosphere that is heated from below only. By this we mean an atmosphere for which the source of radiation is from the interior or ground only. A good example would be a giant planet atmosphere dominated by interior flux. In that case $T_{\rm int} = T_{\rm eff,p}$, where $T_{\rm eff,p}$ is the effective temperature of the giant planet. Another example is a planet with a thin atmosphere that is transparent for the radiation emitted by its host star. Mars is such a case. The solar radiation is absorbed by and heats the Martian surface and is re-emitted at longer infrared wavelengths. Terrestrial planets atmospheres are typically not transparant at IR wavelengths where molecules can absorb and re-emit radiation. This way the atmosphere is heated as radiation travels out to space. In such a scenario, $T_{\rm int} = T_{\rm s}$, where $T_{\rm s}$ is the surface temperature.

Following exactly the same reasoning as in section (10.1), we find

$$T^4(\tau) = \frac{3}{4} T_{\rm int}^4 (\tau + \frac{2}{3}). \tag{12.14}$$

Notice that at the top of the atmosphere ($\tau = 0$), we have $T(0) = 2^{1/4}\ T_{\rm int}$. This matches our result for the simple greenhouse model Eq. (12.9), identifying the surface temperature with $T_{\rm int}$ and the equilibrium temperature with the temperature at the top of the atmosphere.

The emergent total specific intensity is given by Eq. 10.14, with $H = \sigma T_{\rm int}^4/4\pi$. The total emergent flux is $\sigma T_{\rm int}^4 = \sigma T_{\rm eff,p}^4$.

12.2.3 Atmosphere heated from above and below

We now turn to the case of an atmosphere heated from above and below. Again, the aim is to develop an analytic expression for the temperature structure. In constructing this model we want to include a crucial feature of heated planetary atmospheres, namely that the wavelengths of the incoming stellar photons (typically visual light) are different from those of the re-radiated photons (infrared light). In these two wavelength regimes, opacities and hence optical depths may be very different. Moreover, the altitudes at which radiation is absorbed and emitted are different.

So, let us introduce two characteristic wavelength regimes. Wavelength regime 1 for visual light and 2 for infrared light. Strictly speaking, our model is now no longer grey (as we have two wavelengths), but as we will treat this two-opacity model further along similar lines much of the grey nature of the problem is preserved. As two energy flows are considered, it is also referred to as a *two-stream model*.

(Exo)planet atmospheres are by definition cool relative to the temperature of their host stars – so cool that their emission is concentrated at infrared wavelengths. Therefore, we assume that the planetary atmosphere exclusively emits light at infrared wavelength, i.e. the emission of the planet at visual wavelength $\eta_1 = 0$. So, radiation is absorbed at visual and infrared wavelengths, and in this framework only emitted in infrared wavelengths. Consequently,

$$S_1 = \eta_1/\chi_1 = 0. \tag{12.15}$$

Radiative equilibrium constraint Eq. 9.42 (see also Eq. 10.3) now implies that

$$\begin{aligned} \int_1 \chi_\nu \, S_\nu \, d\nu + \int_2 \chi_\nu \, S_\nu \, d\nu &= \int_1 \chi_\nu \, J_\nu \, d\nu + \int_2 \chi_\nu \, J_\nu \, d\nu \\ \chi_1 \int_1 S_\nu \, d\nu + \chi_2 \int_2 S_\nu \, d\nu &= \chi_1 \int_1 J_\nu \, d\nu + \chi_2 \int_2 J_\nu \, d\nu \\ \chi_2 S_2 &= \chi_1 J_1 + \chi_2 J_2, \end{aligned} \tag{12.16}$$

where χ_i, S_i, and J_i are appropriate (semi-total) opacities, source functions, and mean intensities. Similar semi-total quantities will be introduced below without formally defining them. The source function for wavelength regime 2 is then given by

$$S_2 = J_2 + \frac{\chi_1}{\chi_2} J_1 = J_2 + \gamma J_1, \tag{12.17}$$

where the ratio of the visible and infrared mean absorption coefficient γ is assumed constant throughout the atmosphere. The value of γ links the optical depths in the visual and infrared, as

$$d\tau_1 = -\chi_1 \, dz = -\frac{\chi_1}{\chi_2} \, \chi_2 \, dz = \gamma \, d\tau_2. \tag{12.18}$$

The two radiative transfer equations are given by

$$\mu \frac{\mathrm{d}I_1}{\mathrm{d}\tau_1} = I_1 \tag{12.19}$$

$$\mu \frac{\mathrm{d}I_2}{\mathrm{d}\tau_2} = I_2 - S_2 = I_2 - J_2 - \gamma J_1 \tag{12.20}$$

The zero-order moment equations (see Eq. 10.4) become

$$\frac{\mathrm{d}H_1}{\mathrm{d}\tau_1} = J_1 \tag{12.21}$$

$$\frac{\mathrm{d}H_2}{\mathrm{d}\tau_2} = J_2 - S_2 = -\gamma J_1 = -\frac{\mathrm{d}H_1}{\mathrm{d}\tau_1}, \tag{12.22}$$

where in the latter equality we have used Eqs. 12.18 and 12.21. This shows that the total flux $H = H_1 + H_2$ is constant, as is required by radiative equilibrium. All flux removed from the incoming beam in the visual re-emerges in the infrared beam.

The first-order moment equation of the infrared radiative transfer equation is (see e.g. Eq. 10.5)

$$3\frac{\mathrm{d}K_2}{\mathrm{d}\tau_2} = \frac{\mathrm{d}J_2}{\mathrm{d}\tau_2} = 3H_2, \tag{12.23}$$

where we have used the Eddington approximation $J_i = 3K_i$. To obtain the temperature profile we focus on infrared wavelengths, which is where all emission occurs. Of course, the infrared equations also depend on the visible wavelength radiation quantities – for instance, through the constraint of radiative equilibrium (Eq. 12.16).

If the star is positioned at $(\mu_\circ, \phi_\circ)$, attenuation of visual light from the incident beam is given by

$$I_1(\tau_2, \mu) = I^\star(0, \mu)\, e^{\gamma\tau_2/\mu} = I^\star\, e^{\gamma\tau_2/\mu}\, \delta\left(\mu + \mu_\circ\right) \delta\left(\phi - \phi_\circ\right), \tag{12.24}$$

where we have used Eq. (12.18) to express the visual optical depth in infrared optical depth, Eq.(12.10) to describe incident light, and have adopted τ_2 as our measure of depth. The semi-total mean intensity is given by (see Eq. 3.9)

$$\begin{aligned} J_1(\tau_2) &= \frac{1}{4\pi}\int_0^{2\pi}\int_{-1}^{0} I^\star\, e^{\gamma\tau_2/\mu} \delta\left(\mu + \mu_\circ\right) \delta\left(\phi - \phi_\circ\right)\, d\mu\, d\phi \\ &= \frac{1}{4\pi} I^\star\, e^{-\gamma\tau_2/\mu_\circ} \qquad (12.25) \\ &= \frac{1}{4\pi}\frac{T_\star^4}{\sigma}\, e^{-\gamma\tau_2/\mu_\circ} \qquad (12.26) \end{aligned}$$

where $\mu_\circ$ is a positive value, see Eq. (12.10). In the final equality, we associate the stellar intensity to a total stellar brightness temperature $T_\star$ (see Eq. 6.43 for the monochromatic equivalent). Integration of Eq.(12.22) yields

$$H_2(\tau_2) = \frac{\mu_\circ}{4\pi}\, I^\star\, e^{-\gamma\tau_2/\mu_\circ} + \text{constant}, \tag{12.27}$$

where the constant of integration follows from realising that (see Eq. 12.13)

$$H_2(0) = H = \frac{\mathcal{F}}{4\pi} = \frac{\mu_\circ\, I^\star}{4\pi} + \frac{\sigma\, T_{\rm int}^4}{4\pi}, \tag{12.28}$$

hence the constant is $\sigma\, T_{\rm int}^4/4\pi$.

With $H_2(\tau_2)$ known, we may integrate Eq. (12.23) to find

$$J_2(\tau_2) = -\frac{3}{4\pi}\frac{\mu_\circ^2}{\gamma}\, I^\star\, e^{-\gamma\tau_2/\mu_\circ} + \frac{3}{4\pi}\sigma\, T_{\rm int}^4\, \tau_2 + \mathrm{A}, \tag{12.29}$$

where A is a constant. Using $J_2 = \sigma T^4/\pi$, we obtain

$$T^4(\tau_2) = \frac{3}{4}\, T_{\rm int}^4\, \tau_2 - \frac{3}{4}\frac{\mu_\circ^2}{\gamma}\frac{I^\star}{\sigma}\, e^{-\gamma\tau_2/\mu_\circ} + \frac{A\pi}{\sigma} \tag{12.30}$$

To obtain the constant A, we follow a procedure that is in principle identical to that used in calculating the temperature structure for an atmosphere heated from below: we write down the integral for the emergent flux, starting from the total flux equivalent of the outward directed flux Eq. 3.21. (See also the convenient form Eq. 4.56 that we could use directly in the atmosphere heated from below case; for our current problem Eddington-Barbier does not hold). For the (total) specific intensity, we use Eq. (4.53), which features the source function – in our case given by Eq. (12.17). Finally, we realise that the total flux is also given by Eq. (12.13). All this can be condensed to the equation

$$\begin{aligned} \mathcal{F}_2 &= 2\pi \int_0^1 \mu \left[\int_0^\infty S_2 \, e^{-\tau_2/\mu} \, \frac{d\tau_2}{\mu} \right] d\mu \\ &= 2\pi \int_0^1 \mu \left[\int_0^\infty (J_2 + \gamma J_1) \, e^{-\tau_2/\mu} \, \frac{d\tau_2}{\mu} \right] d\mu = \sigma \, T_{\rm int}^4 + \mu_\circ \, I^\star. \end{aligned} \quad (12.31)$$

This results in

$$\begin{aligned} T^4(\tau_2) &= \frac{3}{4} T_{\rm int}^4 \left[\tau_2 + \frac{2}{3} \right] + \\ &\quad \mu_\circ T_\star^4 \left[-\frac{3}{4} \frac{\mu_\circ}{\gamma} e^{-\gamma \tau_2/\mu_\circ} + 1 + \frac{3}{2} \left(\frac{\mu_\circ}{\gamma} \right)^2 - \frac{3}{2} \left(\frac{\mu_\circ}{\gamma} \right)^3 \ln \left(1 + \frac{\gamma}{\mu_\circ} \right) \right] \end{aligned} \quad (12.32)$$

Figure 12.3 shows temperature profiles at a range of angles $\mu_\circ$ for the cases $\gamma = 0.1$ and $\gamma = 10$ and assuming $T_\star / T_{\rm int} = 100^{0.25} = 3.16$. For $\gamma = 0.1$ optical extinction is small and stellar light is absorbed at an optical depth that is larger than $\tau_2 = 2/3$, i.e. the depth where we expect the infrared photosphere to be. In the case $\gamma = 10$ the stellar energy is deposited above the infrared photosphere. For the latter model the temperature profile looks pretty similar to the un-irradiated case, although the asymptotic value of the temperature as $\tau_2 \to 0$ is higher the closer one gets to the substellar point (at $\mu_\circ = 1$). In the case $\gamma = 0.1$ we find that the temperature profile has two regions where T is almost constant – the traditional one at low optical depth but also a second one at $\tau_2 > 1$, where the bulk of the incoming flux is absorbed. This second plateau is the result of the superposition of two independent effects. First, the standard solution of radiating the internal flux $\mathcal{F}_{\rm int}$ with no irradiation. Second, the case $T_{\rm int} = 0$, i.e. a zero net energy solution (i.e. $\mathcal{F} = 0$) where the incoming stellar radiation is re-radiated. The temperature profile arising from this second effect consists of a gentle transition between two asymptotic temperatures, a relatively low one at low optical depth and a relatively high one at high optical depth. The difference between these two asymptotes is $\Delta T / T_\star = (3\mu_\circ^2/4\gamma)^{1/4}$, and so is larger for smaller γ. At $\tau_2 > \mu_\circ/\gamma$ the temperature rise of the zero net energy solution quickly dwindles, which is the plateau seen in the the left panel. At even higher optical depth, the normal thermal profile due to transport of internal energy takes over. In principle, a plateau also occurs in the case $\gamma = 10$, however, it can barely be discerned in the figure.

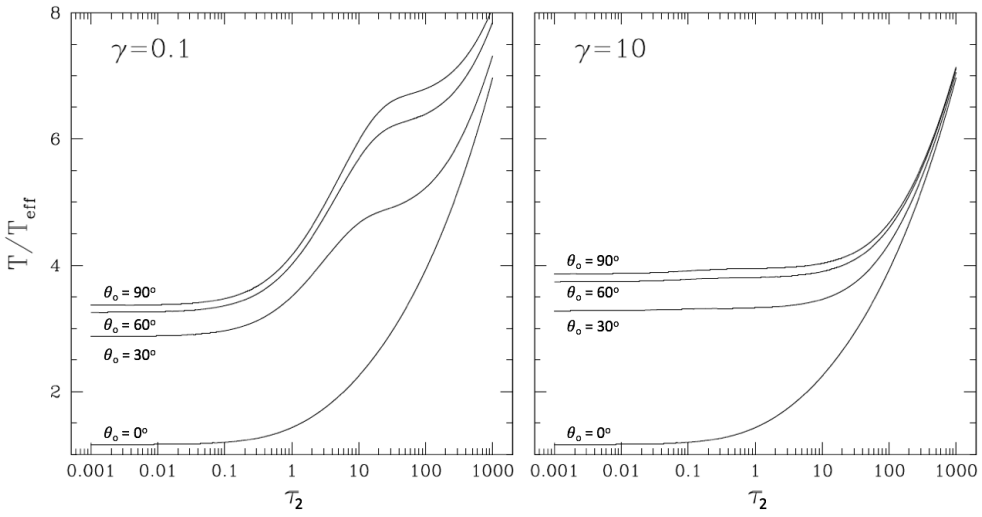


Figure 12.3: *Temperature profiles for the case* $\gamma = 0.1$*, i.e. where the irradiation penetrates to below the re-radiation photosphere, and* $\gamma = 10$*, i.e. where the irradiation is absorbed high in the atmosphere. In all cases* $(T_\star / T_{\rm int})^4 = 100$*. For* $\mu_\circ = \cos\theta_\circ = 0$ *light falls in tangentially (so, we are at the terminator). In that case, the atmosphere is heated from below only as no flux from the host star is absorbed. From: Hansen 2008.*

12.3 Aerosols, clouds, dust, and hazes

The definitions of the four terms in the title of this section is rather ambiguous; see a nice blog by Sarah Hörst on this topic. Following her lead, we define them as follows. Aerosols are any size and kind of particles suspended in a gas. Depending on properties and origin, aerosols can be subdivided in clouds, hazes, and dust[1]. A cloud is a visible mass of liquid and/or solid particles suspended in an atmosphere that form from condensation of atmospheric gases. Cloud material can move back and forth between gas and solid/liquid phase. Cloud droplets prefer to form on pre-existing particles. An example of such pre-existing particles are dust particles, which are solid particles suspended in the atmosphere that did not originate in the atmosphere. On Earth, sources of atmospheric dust are soil particles lifted by weather, volcanic eruptions, and pollution. Haze particles are particles produced from (photo-)chemistry in the atmosphere that results in the formation of in-volatile solids, i.e. that form and persist, unlike cloud material.

The diversity of aerosol materials on different planets is huge, in terms of composition and size- and shape properties, and vertical (and likely lateral) distribution. Earth's clouds are made of liquid water droplets or frozen water crystals ranging from 10 to 100 μm in size. The Venusian atmosphere contains clouds made of sulphuric acid (H_2SO_4). Titan has clouds

[1]Living material, i.e. viruses and bacteria, may be another aerosol category.

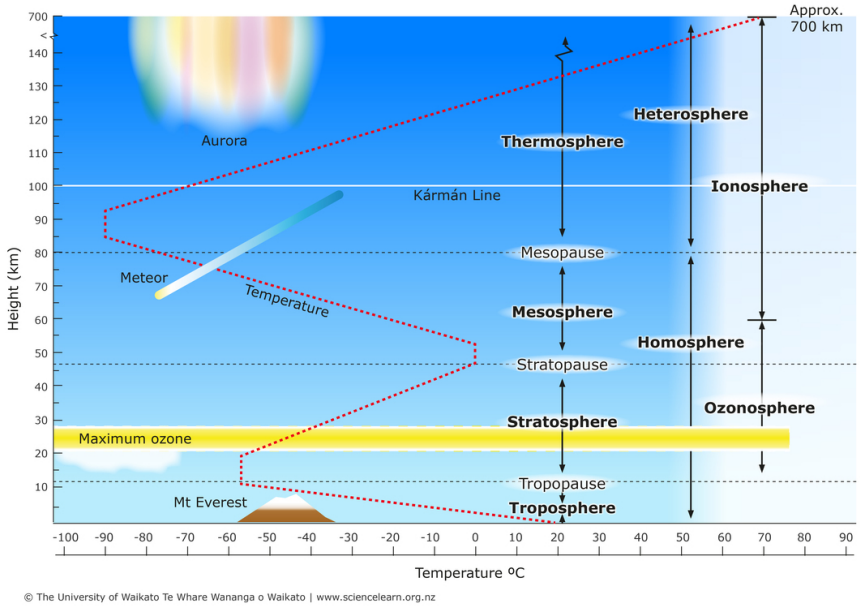


Figure 12.4: *Vertical temperature structure of Earth's atmosphere. From: University of Waikato (2014). For a discussion of the four main layers: troposphere, stratosphere, mesosphere, and thermosphere, see text.*

likely composed of liquid methane. Hot Jupiter-like exoplanets may have clouds composed of liquid iron droplets (as they are far too hot to have water clouds). A single planet can have different cloud layers if its atmosphere spans a temperature range that includes more than one condensible gas.

Clouds contribute to the atmospheric energy balance, e.g. because they may reflect part of the incident light back to space. Some clouds are highly reflective, contributing to a high albedo, such as Earth's water clouds and Venus's sulfuric acid clouds. Hazes on Jupiter and Saturn are strongly absorbing at short wavelengths. Cloud formation is intricately coupled to atmospheric dynamics and temperature. Clouds can block the atmosphere beneath them, weakening the emergent spectral features (see Fig. 2.8). Modelling cloud patterns is extremely challenging.

12.4 The temperature profile of Earth's atmosphere

Figure 12.4 shows the vertical stratification of Earth's atmosphere. It is divided in regions, the lower four corresponding to altitudes of temperature reversals. The *troposphere*, where we live, extends up to an average height of about 9-17 km, depending on latitude (it reaches

highest at the equator). The troposphere is mostly heated through energy transfer from the surface, and its temperature decreases with increasing altitude. The density drops exponentially with height, essentially due to hydrostatic equilibrium (see Sect. 9.3). Half of the total mass of the atmosphere is located in the lower 5.6 km of the troposphere. Nearly all water vapor or moisture is found in the troposphere, hence it is the layer where most of Earth's weather takes place.

The *stratosphere* defines the layer in which temperatures rise with increasing altitude. This phenomenon did not arise naturally from our models of planetary atmospheres and shows that conditions can be more complex than we assumed. The rise in temperature is caused by the local absorption of ultraviolet sunlight by the ozone layer, particularly in the lower stratospheric layers where the overal density is highest and most of the UV radiation is absorbed. At the troposphere-stratosphere boundary or tropopause the temperature may be 210 K, the top of the stratosphere is much warmer, and may be near 0 °C (or 273 K). The stratospheric temperature profile creates very stable atmospheric conditions and lacks the weather-producing air turbulence that is so prevalent in the troposphere.

The *mesosphere* extends from the stratopause at an altitude of about 50 km to the mesopause at 80-85 km. Temperature again drops with increasing altitude due to decreasing absorption of solar radiation by the rarefied atmosphere and increasing cooling by CO_2 line emission. The top of the mesosphere, called the mesopause, is the coldest part of Earth's atmosphere. The mesosphere is the layer where most meteors burn up upon atmospheric entrance.

Above the mesosphere is the *thermosphere*, starting at about 80-85 km and extending to a 500-1000 km height, depending on solar activity. The temperature of the thermosphere gradually increases with height as almost all the solar x-ray and extreme ultraviolet (EUV) at wavelengths < 170 nm is absorbed by the most abundant molecules. Though only a very small fraction of the solar energy is in this spectral range, the density of molecules in the thermosphere is so low that it leads to significant heating. Temperatures are variable, however, and depend on solar activity (e.g. x-ray bursts associated with solar flares; see Sect. 16.4). The low densities cause Kinetic Equilibrium to break down in these layers. The thermosphere is completely cloudless and free of water vapor. The International Space Station orbits in this layer, between 350 and 420 km.

Not shown in Fig. 12.4 is the *exosphere*, which is the outermost layer of Earth's atmosphere. It extends from the exobase, the thermosphere-exosphere boundary at 500-1000 km to about 10 000 km where it merges into the solar wind. The exobase is defined to be the altitude at which the atmosphere becomes collision-less, i.e. where the particle's mean free path (the average distance between particle collisions) is greater than the particle-specific atmospheric scale height.

Convection

Terrestrial planets with thin atmospheres are expected to have convection zones just above the solid surface. The reason is that most of the incident light is absorbed at this surface, and only some in the atmosphere. This causes the surface to be hotter than the overlying atmosphere, i.e. a significant temperature discontinuity arises at the atmosphere-surface boundary. This discontinuity will drive convection. On Earth convection is occurring in most of the troposphere, but this may be different in exoplanets with different atmospheric conditions.

Atmospheric layers with temperature inversions are very stable against convection because the restoring force to a perturbed, lifted air parcel is very strong. As a convective bubble rises, it expands and cools. Being cooler – and hence denser – than the surrounding temperature-inverted atmosphere, the air parcel will sink again.

12.5 Vertical density structure of planetary atmospheres

Planetary atmospheres are – like stellar atmospheres – often assumed to be in hydrostatic equilibrium. The considerations discussed in Sects. 9.3 and 10.2 therefore also apply to planetary atmospheres.

Atmospheric escape

Mercury hardly has an atmosphere, though likely did have so at formation. It appears that over time the inner planet lost its atmosphere to space. Atmospheric escape involves three major stages. First, transport of gases from the lower to the upper atmosphere where escape can take place. Second, conversion from the atmospheric gas (usually in molecular form) to the escaping form (usually atomic or ionic). Third, the actual escape process itself. Any one of these stages can be the bottleneck, that is, the limiting process for atmospheric escape. The actual escape mechanism may differ. Three types of processes are identified: thermal hydrostatic escape, thermal hydrodynamic escape, and non-thermal escape.

Thermal escape or hydrostatic escape. In thermal escape, the thermal velocity exceeds the escape velocity of the planet,

$$\left(\frac{2kT}{m}\right)^{1/2} = \left(\frac{2GM}{R}\right)^{1/2}, \tag{12.33}$$

where the left-hand-side is the most probable velocity of the Maxwell distribution (see Eq. 6.17) for a particles of mass m and the right-hand-side is the surface escape velocity $v_{\rm esc}$, with M the planet mass and R the planet radius. Of course, the Maxwell distribution Eq. 6.16 spreads out over a wide range of velocities, therefore in practice we may multiply the most proba-

ble speed by a factor (larger than unity) to estimate the escape of gas species based on the velocities of the high-velocity tail.

Loss of atmospheric particles by thermal escape occurs in the planetary exosphere, so the relevant temperature in Eq. (12.33) is the temperature in the exosphere.

A more accurate description of escape is to properly account for the Maxwell distribution Eq. 6.16 of the particle being considered. This is known as *Jeans escape*. The Jeans escape flux of particles can be computed by integration over the direction and velocity appropriate for escaping particles

$$\Phi_{\rm Jeans} \;=\; -n_{\rm exob} \int_{v_{\rm esc}(r_{\rm exob})}^{\infty} \int_0^{2\pi} \int_0^{\pi/2} f(v) \cos\theta \sin\theta d\theta d\phi, \qquad (12.34)$$

where the factor $\cos\theta$ assures that the velocity component $v_{\rm z}$ is considered and not the actual value v. In this expression we consider escape to occur from the exobase radius $r_{\rm exob}$, hence $v_{\rm esc}(r_{\rm exob}) = \sqrt{2GM/r_{\rm exob}}$. Integration yields

$$\Phi_{\rm Jeans} \;=\; \frac{n_{\rm exob}}{2\sqrt{\pi}}\, B \left(\frac{2kT_{\rm exob}}{m}\right)^{1/2} (1+\lambda_{\rm exob})\; e^{-\lambda_{\rm exob}}, \qquad (12.35)$$

where $T_{\rm exob} = T(r_{\rm exob})$. The escape parameter

$$\lambda_{\rm exob} \;=\; \frac{GMm}{kT_{\rm exob} r_{\rm exob}} = \frac{r_{\rm exob}}{H} = \frac{E_{\rm esc}}{kT_{\rm exob}}, \qquad (12.36)$$

where the escape energy $E_{\rm esc} = m v_{\rm esc}^2/2$. The factor B accounts for the time lag of repopulation of the high-velocity tail of the Maxwellian distribution and has a value of the order unity ($B \sim 0.5-0.8$). The total mass loss rate of the species being considered is $\dot{M} = 4\pi R^2\, \Phi_{\rm Jeans}$.

Hydrodynamic escape. Hydrodynamic escape occurs when escape is so fast that a planetary wind develops. Actual calculations of hydrodynamic escape are based on the equations of fluid dynamics (see Chapter 9). A hydrodynamic escape state is reached when sufficient energy is deposited in the upper atmosphere of the planet by incident EUV radiation from the host star to drive an outflow. See for instance McCann et al. 2019 (ApJ 873, 89) for such calculations.

Non-thermal escape. In general, a wide variety of processes are included in this type of escape. Generally it refers to collisional processes between charged species that produce atoms energetic enough to escape from a planetary atmosphere to space.

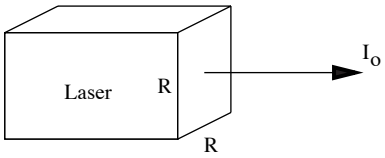
Ablation of the planetary atmosphere by the stellar wind may also be considered a type of non-thermal escape, as well as planetary atmospheres undergoing Roche lobe overflow.

Exercise 12.1

Consider a laser that emits perfectly parallel beams of light

$$I = I_\circ \, \delta(\mu - 1) \tag{12.37}$$

The radiating surface of the laser is a square of R times R cm. We observe the laser from a distance d with a telescope that is pointed in the direction of the laser source, is large in comparison to the laser and can spatially resolve the laser.



b) Give the flux at the surface of the laser.

a) Give an expression for the luminosity L of the laser.

We now place the laser at a distance $3d$ of the telescope. For this new situation:

c) Give the flux as measured by the observer.

Exercise 12.2

This could be a nice exam question. Let us construct a slightly more realistic greenhouse model compared to the one discussed in Sect. 12.1. In this so-called 'leaky' greenhouse model (see Fig. 12.5), the single atmospheric layer may be partly optically thin. We assume that a fraction $\alpha = [0, 1]$ of the infrared radiation emitted by the surface is absorbed by the atmospheric layer.

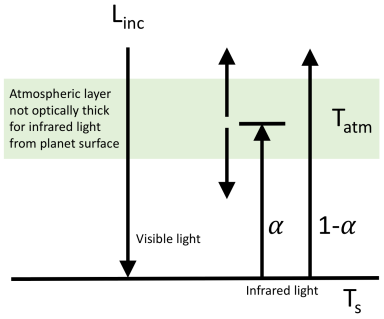


Figure 12.5: *Leaky greenhouse model*

b) Derive the relation between T_s and T_eq.

a) Briefly discuss the limiting cases $\alpha = 0$ and $\alpha = 1$.

c) If the optical depth for IR radiation for the atmospheric layer is τ, how are α and τ related?

Exercise 12.3

Transmission spectroscopy is concerned with measuring the attenuation of beams of stellar light illuminating and (partially) passing through the limb of the planet. Figure 2.8 shows observations using this technique. Plotted on the vertical axis is a measure for the thickness of the atmosphere as a function of wavelength. As this thickness is essentially measured at the planetary limb, it is the vertical thickness relative to some reference value. We are here concerned with deriving an expression that gives this thickness $\Delta R(\lambda)$ (relative to the density scale height H) as function of the measured transit radius $R_{\rm T}(\lambda)$. For convenience, we no longer refer explicitly to wavelength.

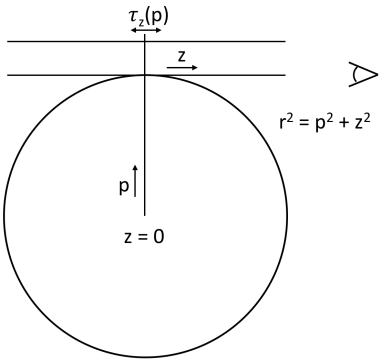


Figure 12.6: *Transient spectroscopy.*

We assume that the equation of state is that of an ideal gas (e.g. Eq. 6.66) and make use of the fact that the absorption occurs high up in the atmosphere, where the temperature profile is essentially isothermal. The temperature of the atmosphere is T, the surface gravity is g, the mean molecular weight μ, and the mass extinction coefficient is $\chi'(\lambda)$. The density scale height is H (see Eq. 9.31).

a) Use the solution for the vertical density structure Eq. 9.30 to show that the optical depth is given by

$$\tau = \tau_\circ \, e^{-r/H}, \tag{12.38}$$

where r is radial height. In this equation, $r_\circ = 0$ is where we calibrate our reference height. At this height $\rho(r_\circ) = \rho_\circ$ and $\tau(r_\circ) = \tau_\circ$. Determine $\tau_\circ$.

b) Give an expression for the optical depth $\tau_\nu(p)$, as function of r, for a beam passing along the limb, i.e. a beam passing through the atmosphere of the planet at fixed impact parameter p. This expression contains an integral; use r as the integration variable.

c) Evaluate the integral obtained in (b). Note that the scale height is small compared to the radial and impact parameter scales for the atmospheric base. As such, the exponential in the integrand will limit significant contributions to scales such that $(r - p) \lesssim H$. Recall that

$$\int_0^\infty x^{-1/2} \, e^{-x} \, dx = \sqrt{\pi}. \tag{12.39}$$

Notice that your result in (c) implies that $\tau_{\rm z}(p = r_\circ)/\tau_\circ = (2\pi r_\circ/H)^{1/2} \gg 1$ because $p/H \gg 1$. This means that the $\tau_{\rm z}(p) = 1$ surface as seen by the observer will lie higher in the atmosphere than $r_\circ$, where $\tau < 1$.

d) We define the transit radius $R_{\rm T}$ as the radius for which $\tau_{\rm z}(p = R_{\rm T}) = 1$. Show that

$$R_{\rm T} = H \ln \left[\left(\frac{2\pi R_{\rm T}}{H} \right)^{1/2} \tau_\circ \right]. \tag{12.40}$$

e) Show that the radial distance ΔR between the transit radius and the radius for which $\tau = \tau_{\rm ref}$ (in the radial direction) is given by

$$\Delta R = \frac{H}{2} \ln \left[\frac{2\pi R_{\rm T}\, \tau_{\rm ref}^2}{H} \right]. \tag{12.41}$$

Like ΔR, the reference optical depth is wavelength dependent. Using models, one may for instance choose it to be the value which corresponds to the depth $\tau_{\rm Ross} = 2/3$. For the scale height, Sing et al. (2016) adopt $H = H_{\rm eq} = \mathcal{R}\, T_{\rm eq}/\mu g$.

13

Spectral lines

The intensity distribution in the spectral line primarily reflects the run of the source function throughout the stellar atmosphere: in the line core one 'sees' layers that are relatively high up in the atmosphere; in the line wings one sees much deeper (see § 4.6). This coupling exists due to the profile function ϕ_ν, which causes the opacity in the line core to be much higher than in the line wing. In this chapter we will study the line profile in more detail. We first define a means to describe the strength of a spectral line. Next, the processes that shape the spectral line are reviewed. We end with a discussion of the curve-of-growth method.

13.1 Describing the line profile

The most complete description of the spectral line is given by its *profile*. The *relative depression* or *absorption depth* of the profile is strickly speaking defined as

$$d_\lambda \equiv 1 - \frac{I_\lambda}{I_\lambda^{\rm c}} \tag{13.1}$$

where $I_\lambda^{\rm c}$ is the continuum intensity at the wavelength λ. In the presence of a spectral line it is by definition impossible to measure the continuum intensity, therefore its value is determined by interpolation of the continuum intensity at both sides of the line profile. In case the absorption depth is positive, the line is an absorption line; in case d_λ is negative, we are dealing with an emission line.

As for almost all stars one can only measure the flux – and not the specific intensity as a function of position on the stellar disk – the absorption depth is often described by

$$D_\lambda \equiv 1 - \frac{\mathcal{F}_\lambda}{\mathcal{F}_\lambda^{\rm c}} \tag{13.2}$$

where $\mathcal{F}_\lambda^{\rm c}$ is the continuum flux at wavelength λ. Also the continuum flux at a wavelength in the line profile can not be measured directly, and therefore it also follows from interpolation of the continuum flux at both sides of the line.

Before the absorption profile can be used as a diagnostic of the stellar atmosphere it needs to be corrected for instrumental distortions. This distortion, which always leads to degradation of the line profile, is described by the *instrumental profile*. Say a light source emits an emission line of infinitesimal width. The profile of this line can be described by a δ-function. Measurement of this line by an instrument shows a smeared profile (typically a Gaussian profile), of which the sharpness is determined by the quality and/or settings of the spectrograph. A measure of this sharpness is the *spectral resolution* or *resolving power*

$$R \equiv \lambda/\Delta\lambda \tag{13.3}$$

Here $\Delta\lambda$ is the bin-width of the flux measurement. For excellent instruments R can be larger than 100 000.

Equivalent width

The concept of *equivalent width* of a spectral line was developed by Marcel Minnaert (1893-1970)[1]. The equivalent width is the line profile integrated absorption depth, i.e.

$$W_\lambda(\mathrm{line}) \equiv \int_{\mathrm{line}} D_\lambda \, d\lambda = \int_{\mathrm{line}} \left(1 - \frac{\mathcal{F}_\lambda}{\mathcal{F}_\lambda^{\mathrm{c}}}\right) d\lambda \tag{13.4}$$

and is – in case of an absorption line – equivalent to the width of a fully blackened rectangular profile of identical surface area (see figure 13.1). Though the cgs unit of equivalent width in the above formula is cm, it is custom to measure W_λ in (milli-) Angström or in the velocity unit $\mathrm{km\,sec^{-1}}$. In case of an absorption line $\mathcal{F}_\lambda^{\mathrm{c}} \times W_\lambda$ is equal to the total continuum energy that is removed by the line. In case of an emission line W_λ will be negative, and $-\mathcal{F}_\lambda^{\mathrm{c}} \times W_\lambda$ describes the total energy that is added by the line to the continuum. The equivalent width is a suited measure for the strength of the spectral line, as it is, for instance, much less sensitive to smearing of the profile as a result of the finite resolution of the spectrograph, than is, for instance, the central absorption depth. For an accurate measurement of W_λ it usually suffices to have a spectral resolution of $R \sim 8\,000$ if the signal-to-noise ratio is at least several tens[2] . The equivalent width offers a quantitative measure of the line profile in cases where the flux levels are too low to observe a detailed line profile. Finally, the equivalent width is independent of interstellar extinction (see § 19.2).

If the source is spatially resolved, one can determine the equivalent width from

$$w_\lambda(\mathrm{line}) \equiv \int_{\mathrm{line}} d_\lambda \, d\lambda \tag{13.5}$$

The value of $w_\lambda(\mathrm{line})$ may vary over the (projected) surface of the source being studied.

[1] Born in Ghent, Belgium, Minnaert worked at the Utrecht astronomical observatory Sonnenborgh from shortly after World War I. From 1937 until 1962 he was the director of the observatory.

[2] See Vollmann & Eversberg 2006 for a discussion on how to estimate the error in equivalent width

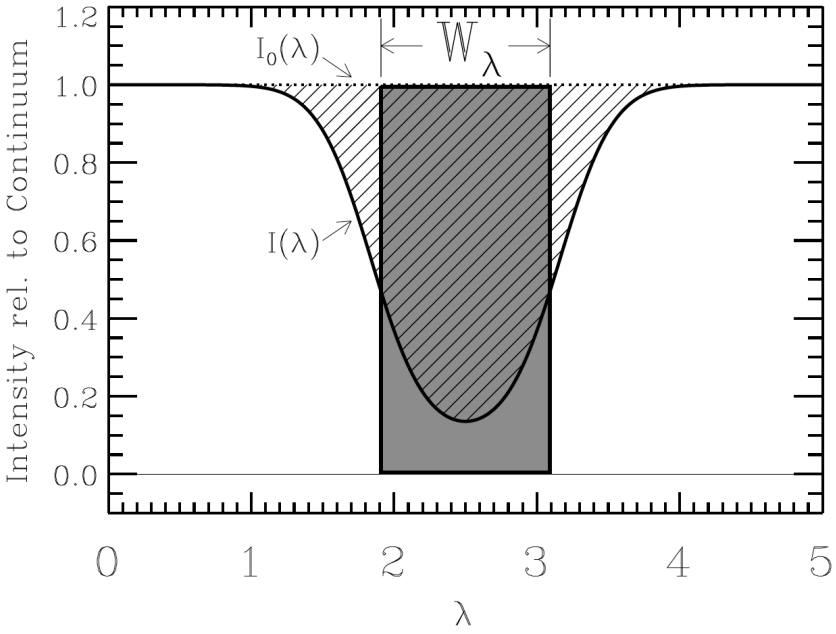


Figure 13.1: *Schematic representation of the flux behaviour of a spectral line (i.e. the profile). The equivalent width W_λ is equal to the surface area of the spectral line divided by the continuum flux. Figure: Edward Jenkins.*

Total line flux

The line profile integrated flux or *total line flux* in erg cm^{-2} sec^{-1} is

$$\mathcal{F}(\text{line}) \equiv \int_{\text{line}} (\mathcal{F}_\nu - \mathcal{F}_\nu^{\text{c}})\, d\nu = \int_{\text{line}} (\mathcal{F}_\lambda - \mathcal{F}_\lambda^{\text{c}})\, d\lambda \tag{13.6}$$

The total line flux (measured at distance d) is especially important in studies of emission lines, and therefore is defined such that a positive value results if $\mathcal{F}_\nu > \mathcal{F}_\nu^{\text{c}}$. One obtains $\mathcal{F}(\text{line}) = -\mathcal{F}_\lambda^{\text{c}} \times W_\lambda = -\mathcal{F}_\nu^{\text{c}} \times W_\nu$. In, for instance, planetary nebulae the continuum flux can be so low that it can not be measured (accurately), i.e. $\mathcal{F}_\nu \gg \mathcal{F}_\nu^{\text{c}}$. The continuum contribution can then simply be omitted when computing the total line flux. Note that the total line flux – in contrast with the equivalent width – is dependent on the amount of interstellar extinction.

Related to the total line flux is the *luminosity in the line*, in erg sec^{-1}, i.e.

$$L(\text{line}) = 4\pi d^2\, \mathcal{F}(\text{line}) \tag{13.7}$$

13.2 Line broadening

There are several processes that can influence the shape of the line profile. We first briefly list these processes, then we discuss each of them seperately:

- natural broadening or radiation damping, due to the finite lifetime of excited states
- Doppler broadening due to the thermal motions of particles
- collisional or pressure broadening, due to the physical interaction of atoms and ions
- broadening due to the rotation of the star
- broadening due to the presence of a stellar wind

Natural broadening

Spontaneous de-excitations cause the lifetime of an excited state to be statistically distributed around a finite value. The Heisenberg uncertainty principle implies that the uncertainty in lifetime Δt must be coupled to an uncertainty in the energy ΔE of the transition, such that $\Delta E \cdot \Delta t \geq h/4\pi = \hbar/2$. In other words, one expects a broadening of the line, which is refered to as *natural broadening* or *radiation damping*. The term "natural" implies that this broadening occurs even if no other particles are present; the term "damping" results from the classical description of the atom as a driven-damped classical oscillator. It is custom to correct the classical result for quantum mechanical effects using the oscillator strength (see § 7.3). Without its formal derivation we give that the profile that belongs to natural broadening is a *Lorentz profile*, i.e.

$$\phi(\nu - \nu_{lu}) = \phi(\Delta\nu) = \frac{1}{\pi} \frac{\gamma^{\rm cl}/4\pi}{(\nu - \nu_{lu})^2 + (\gamma^{\rm cl}/4\pi)^2} = \frac{1}{\pi} \frac{\Delta\nu_{\rm L}}{\Delta\nu^2 + (\Delta\nu_{\rm L})^2} \tag{13.8}$$

where

$$\gamma^{\rm cl} = \frac{8\pi^2 e^2 \nu_{lu}^2}{3 m_{\rm e} c^3} = \frac{8\pi^2 e^2}{3 m_{\rm e} c \lambda_{lu}^2} \tag{13.9}$$

is the *classical damping constant* (dimensions $\rm s^{-1}$) and $\Delta\nu_{\rm L}$ is the Lorentz width, defined as the half width at half maximum in hz

$$\Delta\nu_{\rm L} \equiv \frac{\gamma^{\rm cl}}{4\pi} \tag{13.10}$$

Trivially, the full width at half maximum or FWHM = $2\Delta\nu_{\rm L}$. In the line wings the Lorentz profile drops of as $\Delta\nu^{-2}$. Using $\phi(\nu)\, d\nu = \phi(x)\, dx$, we may write the Lorentz profile in its dimensionless form as

$$\phi(x) = \frac{1}{\pi} \frac{1}{x^2 + 1} \tag{13.11}$$

where $x = \Delta\nu/\Delta\nu_{\rm L}$. The Lorentz profile fulfils the normalization requirement eq. (7.4). Note that the characteristic natural line width in wavelength units, i.e. $\Delta\lambda_{\rm L} = \lambda_{lu}^2\gamma^{\rm cl}/4\pi c$ cm^{-1}, is constant and given by $6\,10^{-5}$ Å. This is a very small value. Usually, other line broadening mechanisms are more important.

The quantum mechanical equivalent of the damping constant is given by

$$\Gamma = \Gamma_u + \Gamma_l = \sum_{i<u} A_{ui} + \sum_{i<l} A_{li} \tag{13.12}$$

It is the sum of all possible spontaneous de-excitations from the upper level plus those from the lower level (per second). Also the lower level plays a role because the width of the spectral line is determined by the uncertainty in energy of both levels. One may simply add the contributions of the upper and lower level as a convolution of two Lorentz profiles with half width Γ_u and Γ_l yields again a Lorentz profile of which the half width is $\Gamma_u + \Gamma_l$. To get an impression of how the classical values of the half width compare to the quantum mechanical values: for the strong Hα-line of hydrogen we find $\Delta\lambda_{\rm L} = 6.5\,10^{-4}$ Å or 0.03 km sec^{-1}. This is about an order of magnitude larger than the classical value. Weak lines, with smaller values of the Einstein coefficients A_{ul} will have smaller natural line widths. For an idealized two level atom it holds that $\Gamma = A_{ul}$.

Doppler broadening

The motion of a radiating particle along the line of sight will cause a Doppler shift given by

$$\frac{\Delta\nu}{\nu} = -\frac{\Delta\lambda}{\lambda} = \frac{\xi}{c} \tag{13.13}$$

Here ξ is the velocity component in the line of sight. A photon of frequency ν in the reference frame of the observer and that is moving toward the observer, is seen to have a frequency

$$\nu' = \nu\left(1 - \frac{\xi}{c}\right) \simeq \nu - \nu_{lu}\frac{\xi}{c} \tag{13.14}$$

by the moving particle. It can therefore only absorb this photon if frequency ν is such that ν' corresponds to a line frequency in the rest frame of the particle. For a minute think of the profile function as a delta function at $\nu' = \nu_{lu}$. Is the particle moving towards us (in which case ξ is positive), the observers frame frequency ν needs to be bluer than ν_{lu} for the particle to be able to absorb it. Likewise, for a particle that is moving away from us (in which case ξ is negative), ν needs to be redder than ν_{lu}.

The distribution of thermal velocities in the line of sight is given by the one dimensional Maxwellian velocity distribution (see eq. 6.14)

$$\frac{n(\xi)}{N}\,d\xi = \frac{1}{\xi_{\rm D}\sqrt{\pi}}\exp\left[-\left(\xi/\xi_{\rm D}\right)^2\right]\,d\xi \tag{13.15}$$

where

$$\xi_{\rm D} = \left(\frac{2kT}{m}\right)^{1/2} = 12.895 \left(\frac{T}{10^4 A}\right)^{1/2} \quad {\rm km\ s^{-1}} \tag{13.16}$$

is the *Doppler velocity*, with m the mass and A the atomic weight (in amu) of the particle.

If thermal Doppler broadening is the only source of line broadening, such that in the co-moving frame the profile function may be replaced by a delta function $\delta(\nu' - \nu_{lu})$ – as in this case the energy levels of the transition are infinitely sharp and the extinction will occur only at a single (shifted) frequency – then one finds for the profile function in the reference frame of the observer

$$\begin{aligned} \phi(\nu - \nu_{lu}) &= \int_{-\infty}^{+\infty} \delta\left(\nu' - \nu_{lu}\right) \frac{n(\xi)}{N}\, d\xi = \int_{-\infty}^{+\infty} \delta\left(\nu - \nu_{lu}\frac{\xi}{c} - \nu_{lu}\right) \frac{n(\xi)}{N} \frac{{\rm d}\xi}{{\rm d}\nu}\, d\nu \\ &= \frac{n((\nu - \nu_{lu})c/\nu_{lu})}{N} \frac{c}{\nu_{lu}} \end{aligned} \tag{13.17}$$

where in the second right hand side we have switched to the usual unit of the profile function (i.e. ${\rm hz}^{-1}$) and in the third right hand side we have taken $\xi = (\nu - \nu_{lu})c/\nu_{lu}$, such that $d\xi/d\nu = c/\nu_{lu}$. One finds

$$\phi(\nu - \nu_{lu}) = \phi(\Delta\nu) = \frac{1}{\sqrt{\pi}\Delta\nu_{\rm D}} \exp\left[-\left(\Delta\nu/\Delta\nu_{\rm D}\right)^2\right] \tag{13.18}$$

where

$$\Delta\nu_{\rm D} \equiv \frac{\xi_{\rm D}}{c}\nu_{lu} = \frac{\nu_{lu}}{c}\left(\frac{2kT}{m}\right)^{1/2} \tag{13.19}$$

is the *Doppler width*. Equation (13.18) is referred to as the *Doppler profile*. The FWHM of this profile is $2\sqrt{\ln 2}\,\Delta\nu_{\rm D}$. In a stellar spectrum one observes the combined effect of a large number of atoms, that each have their own thermal velocity projected along the line of sight. As $\xi_{\rm D}$ is proportional to $A^{-1/2}$ large differences in the characteristic width of thermal profile functions may exist. For a star like the sun ($T \sim 6000$ K) the Doppler width for hydrogen is typically 10 ${\rm km\,sec^{-1}}$, while for iron it is only 1.3 ${\rm km\,sec^{-1}}$ – more than seven times as small.

The dimensionless Doppler profile function follows after introducing $x = \Delta\nu/\Delta\nu_{\rm D}$

$$\phi(x) = \frac{1}{\sqrt{\pi}} \exp(-x^2) \tag{13.20}$$

This function is properly normalized to unity.

Collisional broadening

Radiating or absorbing atoms and ions in a gas can not be treated as strictly independent particles. The densities are usually so high that the atoms and ions will "feel" other particles,

Table 13.1: *Overview of collisional broadening mechanisms, and an indication of the type of stars for which these are relevant (adapted from Rob Rutten).*

p	Mechanism	Profile	Atom/Ion	Perturbers	Spectral Types
2	linear Stark	Holtzmark	H, hydrogenic	ions	early-type
		Lorentz?	H, hydrogenic	electrons	early-type
3	resonance	Lorentz	atom A (read: H)	atom A (read: H)	solar
4	quadratic Stark	Lorentz	non-hydrogenic	electrons, ions	early-type
6	van der Waals	Lorentz	atom B	atom A (read: H)	late-type

even if this does not result in an extinction or emission of a photon (such as e.g. in a free-free interaction). These other particles can be electrons, ions or atoms of the same kind, or also, in case of cool stars, molecules. In the simplest representation, the energy levels of the atom or ion that absorbs a photon will be somewhat disturbed by Coulomb interaction with other nearby particles, such that (temporarily) their energy will be modified. The extend by which these energy levels will be affected is a function of distance to, and the amount and nature of the perturbing particles. The perturbation of the energy levels leads to a line broadening. This process is refered to as *collisional broadening* or *pressure broadening*.

In many cases the perturbation of the level energies as a function of distance to the particles that are responsible for it can be approximated by a powerlaw, i.e. $\Delta E \propto r^{-p}$ where p is an integer number which depends on the type of interaction. One may expect that the upper level of the transitions is more severely affected by the interaction than is the lower level. The perturbation of the spectral line is simply the difference between the two independent levels, i.e.

$$\Delta E_u - \Delta E_l = h\Delta\nu = \frac{C_p(l,u)}{r(t)^p} \tag{13.21}$$

The interaction constant C_p needs to be measured or computed for each transition and type of interaction (see table 13.1).

The lowest order broadening, $p = 2$, is the *linear Stark effect*. It is particularly important for neutral hydrogen and causes the broad line wings of the Balmer lines, such as Hα, Hβ, Hγ and Hδ, in main sequence stars (see § 2.1). Also hydrogen-like lines, such as the He II and Rydberg lines – i.e. lines between levels with high main quantum numbers – are sensitive for the linear Stark effect. The nature of the interaction is based on the fact that neutral hydrogen has a permanent dipole character as the electron is not capable of fully shielding the proton. The perturbing particles are protons and electrons. The decline of the interaction force is the same as that of the Coulomb field of a point source, which implies that the perturbation of the level energies is dependent on the field strength in the ambient medium.

Resonance broadening describes the mutual interaction between neutral hydrogen atoms, and is based on the permanent dipole nature of these atoms. For this broadening mechanism the decay is described by $p = 3$. The effect plays a role in those stars in which neutral hydrogen

itself is the main perturbing particle. This is the case for solar type stars, where it is relevant for the broadening of the Hα line.

Lines of atoms and ions that are not hydrogen-like, and that therefore do not have a dipole moment, feel a *quadratic Stark effect* because of interactions with protons and electrons. For this effect $p = 4$. In early-type stars interactions of non-hydrogenic atoms and ions with electrons are the dominant cause of collisional broadening.

Interactions of non-hydrogen like atoms with neutral hydrogen are the dominant cause of collisional broadening in solar-type or even cooler stars, brown dwarfs and gas giant planets. This type of collisional broadening is called *Van der Waals broadening* and has $p = 6$.

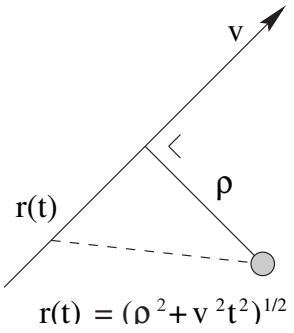


Figure 13.2: *Schematic representation of the collision parameter* ρ

It is beyond the scope of these lecture notes to give a detailed description of the theory of collisional broadening. We mention that there are formalisms to describe the problem in two limiting cases. In the *impact approximation* the duration of the perturbation is short compared to the time between collisions. It is a meaningful way of describing perturbations that are effective over a short distance (i.e. the type of interactions that have high values of p). The result of the perturbation is described as a phase jump in the wave train emitted by the exciting or de-exciting atom. This Lindholm approximation causes a broadening that is described by a Lorentz profile – similar to natural broadening – with a damping constant Γ_p that may be added to eq. (13.12). The effective duration of the perturbation is typically of the order $\Delta t_s \sim \rho_{\rm eff}/v$, where $\rho_{\rm eff}$ is the effective collision parameter (see figure 13.2) and v is the velocity of the perturbing particle. The time interval between two collisions is given by $\Delta t_b = (N\pi\rho_{\rm eff}^2 v)^{-1}$ where N is the density of the perturbers; v the velocity of these particles, and $\pi\rho_{\rm eff}^2$ the effective collision cross section of the particle that suffers the collision. The cross section depends on the type of interaction potential (see eq.13.21). The validity of the impact approximation therefore also depends on the density of the impacting particles. Only when this is sufficiently low, the approximation may be applied, as expected.

For increasing density the interaction time Δt_s will become shorter and shorter. At some point the particles that are being perturbed will feel a continuous interaction due to (partly) overlapping collisions. This allows for a *statistical approximation*, which is the other extreme of describing the collisional broadening problem. The linear Stark effect is described in this way because, due to the relatively slow decay of the interaction potential ($\Delta E \propto r^{-2}$), the effective cross section is very large. In this approximation the profile function will deviate from the Lorentz profile and has a so-called Holtzmark shape. In the line wings this profile decays as $\Delta\nu^{-5/2}$, whereas the Lorentz profile shows a $\Delta\nu^{-2}$ dependence.

Turbulent broadening

The profiles of spectral lines are often broader than one would expect on the basis of the known broadening mechanisms. Examples of physical processes that might cause this broadening are (several types of) waves and turbulent convection. One collectively refers to the motion fields caused by these effects as *turbulence*. We implicitly ignore these turbulent motions when we assume that the stellar atmosphere is time-independent and consists of homogeneous layers. To "compensate" for this simplification we use two *ad-hoc* fit parameters, first introduced by Otto Struve (1897–1963). These are the *microturbulent velocity* and the *macroturbulent velocity*.

It is assumed that the microturbulent motions are random in nature and that they show a Gaussian velocity distribution and correspond to characteristic length scales that are significantly smaller than the geometrical thickness of the line-forming layer in the atmosphere. The effect of microturbulence is a broadening of the profile function. Convolution with the thermal velocity distribution (which truly is Gaussian in nature) yields a new Gauss distribution with a Doppler width

$$\Delta\nu_{\rm D} \equiv \frac{\nu_{lu}}{c}\sqrt{\xi_{\rm D}^2 + \xi_{\rm micro}^2} \tag{13.22}$$

where $\xi_{\rm micro}$ is the microturbulent velocity. Micro turbulence may cause an increase in the equivalent line width. If "saturation" occurs at line centre – i.e. all fotons that can be absorbed at line centre are indeed absorbed – the presence of microturbulent velocity fields will have the effect that the available wavelength regime over which fotons can be absorbed is increased. In other words, more particles will be able to absorb line (wing) fotons, such that W_λ can increase further (see section 13.4).

Also the macroturbulent velocity distribution is often assumed Gaussian. Here, however, one assumes that the characteristic length scales of the gas elements that experience macroturbulence is larger than the geometrical thickness of the line-forming layer. If one could observe the stellar disk with sufficient spatial resolution, the macroturbulence would manifest itself as small "wiggles" on the spectral lines. During excellent atmospheric conditions this has indeed been observed for the sun (one finds velocity amplitudes of ~ 0.3 to $\sim 3\ \mathrm{km\,sec^{-1}}$). The surface integrated effect of macroturbulence is that the line profiles become broader and less deep; the equivalent line width of the spectral line is not affected. The assumption of a Gaussian distribution of the moving elements responsible for the macroturbulent motion field is only justified when the velocity fluctuations are not coupled to temperature fluctuations, such as could occur if rising gas elements have a higher temperature relative to those that sink (think about granulation on the solar surface). If this is the case, then one should apply a "multiple-stream model" for which asymmetric line profiles may occur.

Macroturbulence is (most simplistically) accounted for by convolving the computed absorption depth with a Gaussian velocity distribution

$$D_\lambda = D_\lambda^{\rm comp} * \frac{1}{\xi_{\rm macro}\sqrt{\pi}}\exp(-\xi^2/\xi_{\rm macro}^2) \tag{13.23}$$

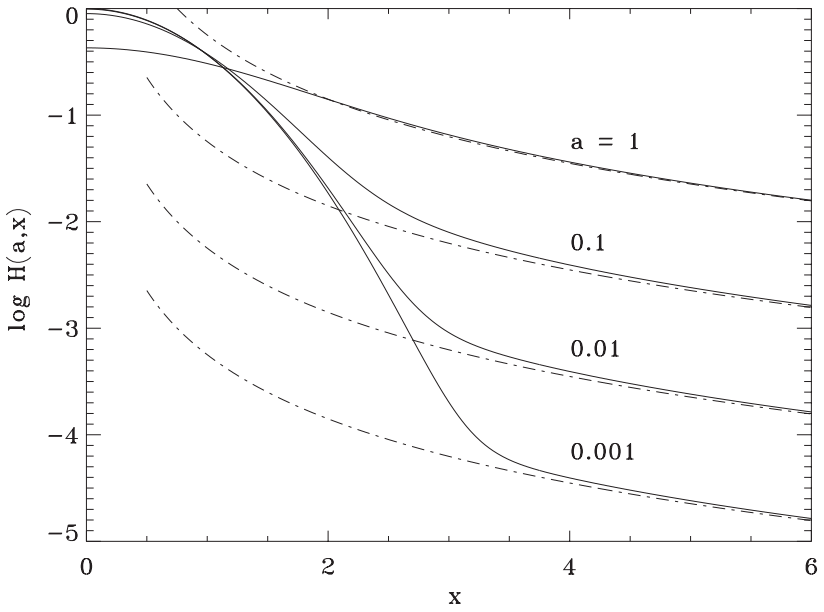


Figure 13.3: *The Hjerting function $H(a,x)$ as function of x for different values of the damping parameter a. The dashed line gives the approximation $H(a,x) \simeq a/\sqrt{\pi}x^2$ for the damping wings.*

Voigt profile function

If the collisional broadening may be described by a Lorentz profile (such as in the impact approximation) the total damping profile will again be a Lorentz profile, with damping constant $\gamma = \Gamma + \Gamma_p$. Assuming the collisional broadening and the thermal broadening are independent processes, the final profile function will be a convolution of a Lorentz and a Doppler profile. As the damping profile is properly normalized we may replace the δ-function in eq. (13.17) by a Lorentz profile. This yields the *Voigt profile*

$$\begin{aligned}
\phi(\nu-\nu_{lu}) &= \left[\frac{1}{\sqrt{\pi}\Delta\nu_{\rm D}}\exp\left[-\left(\Delta\nu/\Delta\nu_{\rm D}\right)^2\right]\right] * \left[\frac{\gamma/4\pi^2}{(\nu-\nu_{lu}(\xi/c)-\nu_{lu})^2+(\gamma/4\pi)^2}\right] \\
&= \frac{1}{\sqrt{\pi}\Delta\nu_{\rm D}}\int_{-\infty}^{+\infty}\frac{(\gamma/4\pi^2)\exp\left[-\left(\Delta\nu/\Delta\nu_{\rm D}\right)^2\right]}{(\nu-\nu_{lu}(\xi/c)-\nu_{lu})^2+(\gamma/4\pi)^2}\,d\nu \\
&= \frac{1}{\sqrt{\pi}\Delta\nu_{\rm D}}H(a,x) \qquad (13.24)
\end{aligned}$$

where

$$H(a,x) \equiv \frac{a}{\pi}\int_{-\infty}^{+\infty}\frac{\exp(-y^2)}{(x-y)^2+a^2}\,dy \qquad (13.25)$$

$$y \equiv \frac{\xi}{\xi_{\rm D}} = \frac{\xi}{c}\frac{\nu_{lu}}{\Delta\nu_{\rm D}} \tag{13.26}$$

$$x \equiv \frac{\Delta\nu}{\Delta\nu_{\rm D}} = \frac{\nu-\nu_{lu}}{\Delta\nu_{\rm D}} \tag{13.27}$$

$$a \equiv \frac{\Delta\nu_{\rm L}}{\Delta\nu_{\rm D}} = \frac{\Gamma}{4\pi\Delta\nu_{\rm D}} \tag{13.28}$$

The function $H(a,x)$ is the *Hjerting function*. Examples of Hjerting profiles are given in figure 13.3. The Hjerting profile is not normalized, but has a surface $\sqrt{\pi}$ in units of x. The maximum value at profile centre is about $H(a,x=0) \sim 1-a$ for $a < 1$. In most astrophysical circumstances the value of a is between 10^{-3} and 10^{-1}. For $a \ll 1$ the Hjerting function can be approximated quite accurately by the sum of a Gaussian core and the damping wings of the Lorentz profile, i.e.

$$H(a,x) \simeq \exp(-x^2) + \frac{a}{\sqrt{\pi}x^2} \tag{13.29}$$

For $a \sim 10^{-2}$ the transition between the Doppler core and the Lorentz wings is near $x = 2.7$, where $H(a,x) \sim 10^{-3}$. It shows that the Doppler core extends quite far. Consequently, the fact that in most cases collisional broadening is poorly known is irrelevant; after all the profiles are Doppler shaped until far in the line wings. Only for really strong lines the wings will be sufficiently strong to be insensitive to the Doppler velocity:

$$\phi(\nu-\nu_{lu}) \sim \frac{1}{\sqrt{\pi}\Delta\nu_{\rm D}}\frac{a}{\sqrt{\pi}x^2} = \frac{1}{\sqrt{\pi}\Delta\nu_{\rm D}}\frac{\gamma}{4\pi\Delta\nu_{\rm D}}\frac{\Delta\nu_{\rm D}^2}{\sqrt{\pi}\Delta\nu^2} = \frac{\gamma}{4\pi^2\Delta\nu^2} \tag{13.30}$$

13.3 Rotational broadening

Line broadening through motion of the stellar surface

We first discuss the general theory of line broadening through motion of the stellar surface (following Underhill, 1968). The coordinate system adopted is shown in Fig. 13.4. Here the (x,y)-plane is the plane of the sky and the z-axis points towards the observer. The origin, O, of the coordinate system is at the centre of the star. The position of any point P on the surface of a star is determined by the distance R which is the distance OP, the angle φ which gives the rotation of the plane through the z-axis and the line OP about the z-axis, and the angle θ which is measured in the plane zOP from the z-axis towards the xy-plane. The angle φ varies from 0 to 2π; the angle θ varies from 0 to π/2.

Let ds be an element of surface around the point P perpendicular to the line OP and let ds_z be the projected element of area perpendicular to the z-axis. Let $I_z(\lambda')$ be the specific intensity at wavelength λ' (as seen by someone at P who is co-moving with the rotating star) emergent at the point P in a direction parallel to the z-axis. The flux emitted by the star (in its own

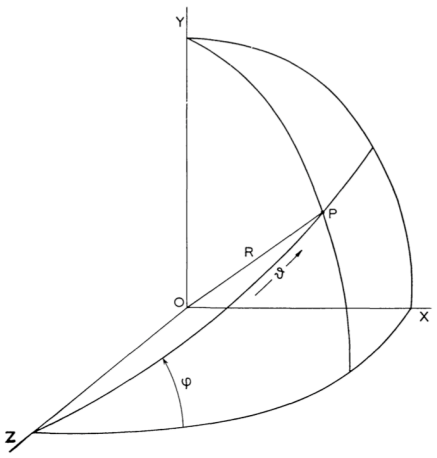


Figure 13.4: *The coordinate system used in describing rotational broadening. From: Underhill (1968).*

reference frame) in the direction of the observer is (see also the description of the ray-by-ray solution in section 3.4)

$$\mathcal{F}_{\lambda'} = \int_S I_z(\lambda')\, ds_z \tag{13.31}$$

where the integration is carried out over the total surface S which faces in the direction of the observer. In general ds_z can be expressed as an analytical function of the angles φ and θ and of a unit distance $R_\circ$. The exact form of the relation depends upon the shape of the star. For a spherical star $R(\theta, \varphi) = R_\circ$, where the latter is the stellar radius. The integration over the surface is performed by letting φ and θ vary over their whole range. If the element of surface at point P has a component of velocity $v_z(P)$ in the z-direction, such that v_z is positive for a surface element moving away from the distant observer, radiation of wavelength

$$\lambda' = \lambda \left(1 - \frac{v_z}{c}\right) \tag{13.32}$$

as seen by a co-moving observer at P, will appear to a distant observer to have the wavelength λ. To this external distant stationary observer the wavelength λ' is a function of the position of point P on the stellar surface. Thus the flux at wavelength λ, as seen by the distant observer, may be found from

$$\mathcal{F}_{\lambda} = \int_S I_z(\lambda')\, ds_z \tag{13.33}$$

The above equation is entirely general. Particular cases such as the change in shape of spectral lines due to rotation of the star, to pulsation or to macro-turbulence can be investigated by substituting in eq. (13.33) particular expressions for $I_z(\lambda')$ and for s_z. In the following the case of the line profile resulting from a rotating star will be investigated.

The particular case of a rigidly rotating star

Consider the case of a spherical star of radius R rotating about the y-axis as a rigid body with an equatorial velocity $v_{\rm R}$. It follows that at point $P(\theta, \varphi)$

$$ds_z = R \sin\theta \, \cos\theta \, d\theta \, d\varphi \tag{13.34}$$

and

$$v_{\rm R}(P) = v_{\rm R} \sin\theta \, \cos\varphi. \tag{13.35}$$

Futhermore one can write with adequate accuracy

$$\lambda' = \lambda - \frac{v_{\rm R}}{c} \lambda_{lu} \sin\theta \, \cos\varphi \tag{13.36}$$

where λ_{lu} is the wavelength at the centre of the line and λ is any neighbouring wavelength such that $\Delta\lambda = \lambda - \lambda_{lu}$ is small with regard to λ_{lu}. The emergent intensity in the direction of the z-axis, I_z at the point P is a function of the angles θ and φ as a result of, for instance, limb darkening or effects of rotation. It also depends on the angles θ and φ through the wavelength λ', see eq. (13.36). In general, even for the case of a spherical star a numerical approach is taken. However, if we assume the intrinsic profile of the line to be infinitely narrow and independent of position on the stellar disk and when the limb darkening in the continuum spectrum may be represented by a linear function of μ (such as the limb darkening law proposed by Milne, see vgl. 11.29) the line broadening may be treated analytically. This has first been shown by Shajn & Struve (1929). Following Unsöld (1955), we introduce the rotational displacement of the line

$$\Delta\lambda = \frac{\lambda_{lu}}{c} \, \omega \, x \sin i \tag{13.37}$$

where the angular velocity $\boldsymbol{\omega}$ is a vector in the (y, z)-plane that is inclined at an angle i from the z-axis towards the observer in the direction of the y-axis. If i = 0°, we see the star pole-on; if it is 90°, we see the star face-on. The angular speed is $\omega = |\boldsymbol{\omega}|$. If we see the star face-on, the rotational velocity at the equator is $v_{\rm R} = \omega R$. If we see the star at an inclination, this maximum rotational velocity is only $v_{\rm R} \sin i$. The maximum velocity displacement is therefore $\Delta\lambda_R = (\lambda_{lu}/c) \, \omega R \sin i = (\lambda_{lu}/c) \, v_{\rm R} \sin i$.

Let us introduce the normalized x and y distance $x' = x/R$ and $y' = y/R$, such that

$$\frac{\Delta\lambda}{\Delta\lambda_R} = x'. \tag{13.38}$$

Following Unsöld, we adopt for the emergent intensity

$$I(\mu) = \alpha \, (1 + \beta \cos\theta) = \alpha \, (1 + \beta\mu) \tag{13.39}$$

where, as usual, θ is the angle between the radial direction and that of the beam towards the observer. The angle μ will be the same for each circle around O, i.e.

$$\sin\theta = \sqrt{x'^2 + y'^2} \tag{13.40}$$

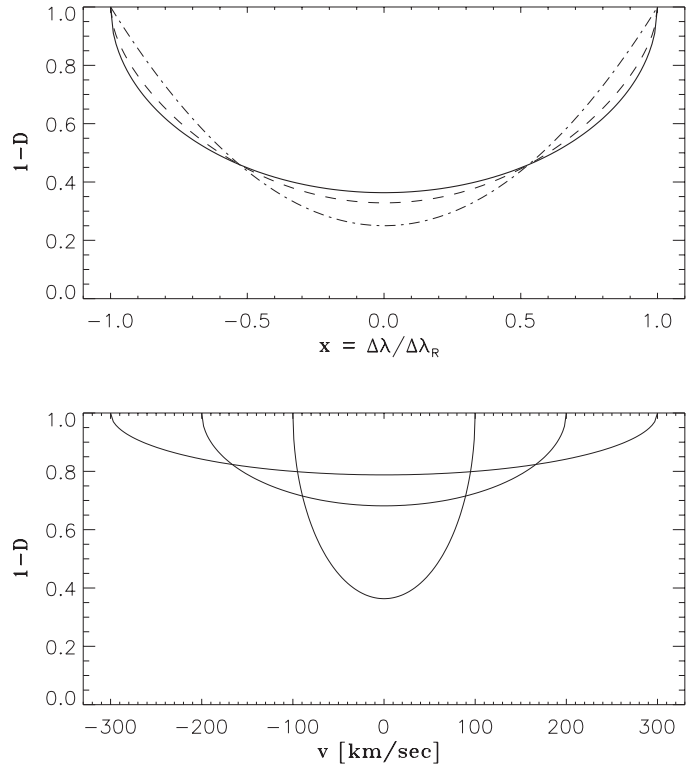


Figure 13.5: *Rotational broadening profiles for a rigidly rotating spherical star assuming the intrinsic profile of the line to be infinitely narrow and independent of position on the stellar disk. The adopted limb darkening law is a linear function of μ, given by eq. (13.39). Plotted is the line flux relative to the continuum, i.e. $1 - D(x')$. Top panel: the horizontal axis shows the wavelength shift $\Delta\lambda$ in units of $\Delta\lambda_R$ (so in units of x'). Profiles for β = 0 (solid line), 3/2 (dashed line) and ∞ (dashed-dotted line) are shown. Bottom panel: the horizontal axis shows the wavelength in velocity units relative to line center. For three values of $v_R \sin i$ (and β fixed to zero) de profile of a line that has an equivalent width in velocity space W_v of 100 $\mathrm{km\,sec^{-1}}$ is shown.*

such that

$$I(\mu) = \alpha \left(1 + \beta\sqrt{1 - (x'^2 + y'^2)}\right). \tag{13.41}$$

If we normalize the flux in the line to unity in x' space, we find for the absorption depth

$$D_{x'} = \frac{\int_0^{\sqrt{1-x'^2}} I(x', y')\, dy'}{\int_{-1}^{+1} \int_0^{\sqrt{1-x'^2}} I(x', y')\, dy'\, dx'} = \frac{\frac{2}{\pi}\sqrt{1 - x'^2} + \frac{\beta}{2}(1 - x'^2)}{1 + \frac{2}{3}\beta} \tag{13.42}$$

For $\beta = 0$, i.e. no limb darkening, the line profile shape is that of an ellips. For the limb darkening law derived for the gray atmosphere (eq. 10.15) the coefficients $\alpha = 2/5$ and $\beta = 3/2$. In figure 13.5 examples of rotationally broadened profiles are shown. The top panel

shows the effect of the limb-darkening coefficient β. In the lower panel $\beta = 0$ is adopted, but profiles are plotted as a function of velocity shift relative to line center. This shows the effect of different values of $v_R \sin i$. Rotation of the stellar surface will not affect the equivalent width of the spectral line as only the observed velocity (frequency) of line flux is changed. The equivalent width in velocity space $W_v = 100\,\mathrm{km\,sec^{-1}}$. Notice that analyzing rapidly rotating stars seen fairly edge on tends to become more challenging as the lines will become shallow and broad, making them more difficult to distinguish from the continuum, especially so if the density of spectral lines in a given frequency interval is large and/or the spectrum has a fairly low signal-to-noise ratio.

13.4 Curve of growth

In the not so distant past, the *curve of growth* analysis, developed by Marcel Minnaert (1934), was a commonly used method to analyze stellar atmospheres as it provides a simple and quick way to estimate a number of important parameters – such as column depths, excitation- and ionization temperatures, and abundances – by only using equivalent line widths. Nowadays the method is not often employed anymore in the analysis of starlight, as thanks to the rapid development in computer technology it has become fairly easy and efficient to use LTE/NLTE model atmospheres. However, the method is still very much used in the study of Quasi Stellar Objects (QSOs) and Gamma Ray Burst (GRB) afterglows (see e.g. Zhang et al. 1997 and Prochaska 2006). We will discuss the curve of growth in detail as the method provides important insight in the physics of the formation of line profiles.

We consider a simple model for a stellar atmosphere consisting of two components: a geometrically deep layer responsible for the continuum, which emits at a brightness temperature T_b; and a homogeneous layer that is located further out, where the spectral line is formed, and that emits according to a Planck function at temperature T_L. We will assume that $T_\mathrm{L} < T_\mathrm{b}$, such that an absorption line is formed. Note that if the temperature in the line forming layer would be higher than T_b the following discussion will remain valid.

The emerging intensity follows from eq. (4.49) and is equal to

$$I_\lambda = B_\lambda(T_\mathrm{b})e^{-\tau_\lambda} + B_\lambda(T_\mathrm{L})\left[1 - e^{-\tau_\lambda}\right] \tag{13.43}$$

where τ_λ is the optical depth in the spectral line. If we neglect the contribution of stimulated emission in the line extinction coefficient, such that eq. (7.32) reduces to

$$\chi_\lambda = \alpha_{lu}(\lambda)n_l = \frac{\pi e^2}{m_\mathrm{e}c}\frac{\lambda_{lu}^2}{c}f_{lu}\frac{1}{\sqrt{\pi}\Delta\lambda_\mathrm{D}}\,n_l\,H(a,x) \tag{13.44}$$

where we have specified the profile function by the Voigt profile eq. (13.24). If we define the integrated *column depth* in $\mathrm{cm^{-2}}$ of particles in level l that are in the line of sight throughout the line forming region as

$$N_l \equiv \int n_l(s)ds \tag{13.45}$$

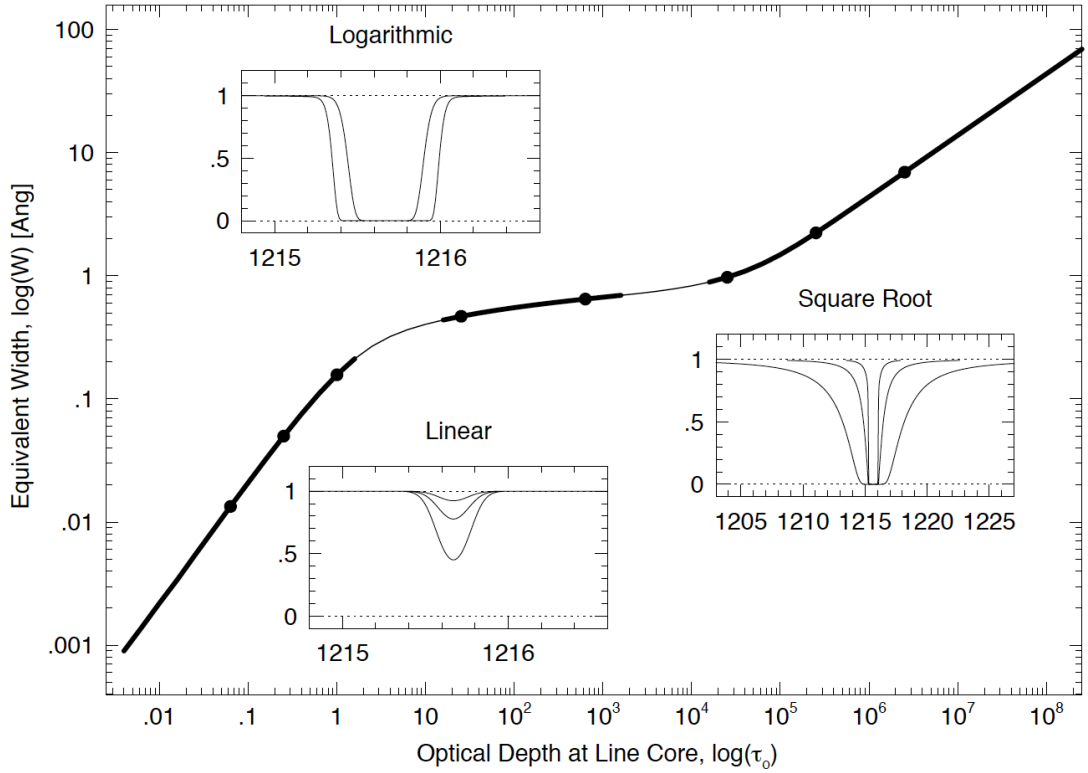


Figure 13.6: *The curve of growth for a Lyα of H I with a Doppler velocity $\xi_{\rm D} = 30\,{\rm km\,sec^{-1}}$. The three regimes discussed in the text, the linear (weak lines), flat (saturated lines), and damping part (strong lines) are shown by thicker curves. The approximations describing these regimes are Eq. 13.49, 13.51, and 13.53, respectively. Corresponding line absorption profiles are shown for each regime and their locations on the COG are marked with filled dots. Figure: Chris Churchill.*

we obtain for the line optical depth

$$\tau_\lambda = \alpha_{lu} N_l = \frac{\pi e^2}{m_{\rm e} c} \frac{\lambda_{lu}^2}{c} f_{lu} \frac{1}{\sqrt{\pi}\Delta\lambda_{\rm D}} N_l\, H(a,x) = \frac{\tau_\circ}{H(a,0)}\, H(a,x) \simeq \tau_\circ\, H(a,x) \quad (13.46)$$

where $\Delta\lambda_{\rm D}$ is the Doppler width, a and x the Voigt parameters as defined in eq. (13.28), and $\tau_\circ$ the optical depth at line centre. Often $a \ll 1$, such that $H(a,0) \simeq 1 - a$, which justifies the last equality. For the relative depression d_λ we get

$$d_\lambda = 1 - \frac{I_\lambda}{I_\lambda^{\rm c}} = \frac{B_\lambda(T_{\rm b}) - B_\lambda(T_{\rm L})}{B_\lambda(T_{\rm b})} \left(1 - e^{-\tau_\lambda}\right) \equiv d_\lambda^{\rm max} \left(1 - e^{-\tau_\lambda}\right) \quad (13.47)$$

with $d_\lambda^{\rm max}$ the maximum depression. This yields for the equivalent line width

$$w_\lambda = d_\lambda^{\rm max} \int_{\rm line} \left(1 - e^{-\tau_\lambda}\right) d\lambda \quad (13.48)$$

Figure 13.6 shows how the line profile and equivalent line width depend on optical depth $\tau_\circ$ for a line that has $d_\lambda^{\max} = 1$, i.e. in case the source function in the line forming layer is negligibly small compared to the continuum source function. The behaviour of the equivalent width may be characterized by three regimes.

Weak lines

First, the regime of weak lines, for which $\tau_\lambda \ll 1$. In this limit the relative depression reduces to $d_\lambda \simeq d_\lambda^{\max}\tau_\lambda$. Say, we can measure the absorption depth of the profile with an accuracy of one promille of the continuum intensity. Figure 13.3 tells us that for a spectral line that has $a = 0.01$ the Doppler core will extend up to $H(a,x) \sim 0.01$. A weak line with a relative depression of 0.1 at line centre would then reach the measurement noise level at the point where the Lorentz wing is about to become dominant. For weak lines we may thus replace the Hjerting function $H(a,x)$ by $\exp[-(\Delta\lambda/\Delta\lambda_\mathrm{D})^2]$, which has surface $\sqrt{\pi}\Delta\lambda_\mathrm{D}$. This yields

$$w_\lambda \simeq d_\lambda^{\max}\tau_\circ\sqrt{\pi}\Delta\lambda_\mathrm{D} = \frac{\pi e^2}{m_\mathrm{e}c}\frac{\lambda_{lu}^2}{c} f_{lu}\, d_\lambda^{\max}\, N_l \tag{13.49}$$

The equivalent width of weak (unsaturated) lines increases linearly with column depth N_l and is independent of the profile function.

Rearranging the above relation yields for the column depth in cm^{-2}

$$N_l\, f_{lu}\, d_\lambda^{\max} = 1.13\times10^{12}\left(\frac{w_\lambda[\mathrm{cm}]}{\lambda_{lu}^2[\mathrm{cm}]}\right) = 1.13\times10^{20}\left(\frac{w_\lambda[\text{Å}]}{\lambda_{lu}^2[\text{Å}]}\right) \tag{13.50}$$

Saturated lines

If $\tau_\circ > 1$, the core of the line becomes saturated. The intensity at line centre approaches the value $B_\lambda(T_\mathrm{L})$ reflecting the maximum depression $d_\lambda^{\max}$. The width of the line wings still increases, however, the corresponding increase in the equivalent width is no longer linear with $\tau_\circ$, but proceeds at a slower pace. It approximately holds that

$$w_\lambda \approx d_\lambda^{\max}\, Q(\tau_\circ)\, \Delta\lambda_\mathrm{D}. \tag{13.51}$$

For increasing optical depth, $Q(\tau_\circ)$ increases from about 2 to 6 in this "shoulder part" of the curve of growth. An approximate analytical expression for this function is $Q(\tau_\circ) = 2\sqrt{\ln\tau_\circ}$, valid from $\tau_\circ \gtrsim 3$. This regime of the curve of growth is therefore (also) often refered to as the "logarithmic part".

Strong lines

For very strong lines, i.e. lines for which $\tau_\circ \gg 1$, the line core is completely saturated and does not react to a further increase of $\tau_\circ$. However, the far line wings will still have $\tau_\lambda < 1$. For

a sufficiently large $\tau_\circ$ both wings will contribute significantly because they are formed in the damping part of the Voigt profile, i.e. where $H(a,x) \simeq a/(\sqrt{\pi}x^2) = (a/\sqrt{\pi})(\Delta\lambda_\mathrm{D}/\Delta\lambda)^2 \sim 1/\Delta\lambda^2$. This decrease with wavelength is much less dramatic than the exponential decay of the Doppler core. In the damping part of $H(a,x)$ we may write

$$\tau_\lambda = \tau_\circ \frac{a}{\sqrt{\pi}x^2} = \tau_\circ \frac{a}{\sqrt{\pi}} \frac{\Delta\lambda_\mathrm{D}^2}{\Delta\lambda^2} \tag{13.52}$$

Using the transformation $u^2 = 1/\tau_\lambda$, we obtain after substitution in eq. (13.48) for the equivalent width

$$\begin{aligned} w_\lambda &\simeq d_\lambda^{\max} \Delta\lambda_\mathrm{D} \sqrt{\tau_\circ \frac{a}{\sqrt{\pi}}} \int_\mathrm{line} (1 - e^{-1/u^2})\, du \\ &= d_\lambda^{\max} \Delta\lambda_\mathrm{D} \sqrt{\tau_\circ a}\, 2\pi^{1/4} \end{aligned} \tag{13.53}$$

For the last equality we used the standard integral

$$\int_{-\infty}^{+\infty} \left(1 - e^{-1/x^2}\right) dx = 2\sqrt{\pi} \tag{13.54}$$

The assumption that we have used here to derive the curve of growth, i.e. an atmosphere consisting of a homogeneous layer that is in LTE, such that $S^\ell = B_\lambda(T_\mathrm{L})$, placed in front of a layer from which a continuum is emitted that is also a Planck field, is called the *Schuster-Schwartzschild* model. It will not come as a surprise that this model does not provide a very realistic representation of the stellar atmosphere. An improvement that can still be done analytically is the *Milne-Eddington approximation*. Here it is assumed that the source function is a linear function of the continuum optical depth, that LTE holds, and that the profile function $\phi(\lambda - \lambda_{lu})$ and the ratio between the line and continuum extiction $\wp_\lambda = \chi_\lambda/\chi_\mathrm{c}$ are independent of depth in the atmosphere. The result $w_\lambda(\tau_\circ)$ turns out to be identical to the Schuster-Schwartzschild result if $\tau_\circ$ is replaced by $\wp_\lambda$. We therefore do not discuss details of this approximation.

Empirical curve of growth

In real life the curve of growth method is applied somewhat different than described above. On the vertical axis one plots $\log W_\lambda/\lambda$. One therefore does not divide by $\Delta\lambda_\mathrm{D} = \lambda\xi_\mathrm{D}/c$, after all the Doppler and microturbulent velocities ξ_D are not known, but by the wavelength, such that we obtain a quantity that is independent of wavelength of the applied set of lines. The horizontal axis, in principle, contains the column depth of the particles responsible for the spectral line, i.e. $N_l = \int n_{ljk}\, ds$. Splitted out in terms of excitation, ionization, and elemental abundance, one obtains for the particle density

$$\begin{aligned} n_{ljk} &= \frac{n_{jlk}}{N_{jk}} \frac{N_{jk}}{N_k} \frac{N_k}{N_N} N_N \\ &= \frac{g_{ljk}}{U_{jk}(T_\mathrm{ext})} e^{-E_{ljk}/kT_\mathrm{ext}} q_{jk}(T_\mathrm{ion}) A_k N_N \end{aligned} \tag{13.55}$$

where we have adopted the LTE formalism. The excitation fraction is described by eq. (6.22) and the ionization fraction q_{jk} by eq. (6.33). $T_{\rm ext}$ is the excitation temperature and $T_{\rm ion}$ the ionization temperature. In LTE these are locally equivalent. However, the curve of growth measures column densities, implying that $T_{\rm ext}$ and $T_{\rm ion}$ represent temperatures that correspond to excitation and ionization fractions that are averaged over this entire column. These mean temperatures need not be equal, therefore they both are introduced.

In many applications of the curve of growth the excitation energy is expressed in electron volt. Using that the temperature associated with 1 eV is equal to $E/k = 11604.5$ K, one finds

$$e^{-E_{ljk}({\rm erg})/kT} = e^{-11604.8\,E_{ljk}({\rm eV})/T} = 10^{-5039.9\,E_{ljk}({\rm eV})/T} \tag{13.56}$$

Taking the energy in eV, the horizontal axis of the curve of growth becomes

$$\log X = \log C + \log(g_{ljk} f_{lu} \lambda_{lu}) - 5039.9\,E_{ljk}/\,T_{\rm ext} \tag{13.57}$$

where

$$C = \frac{\pi e^2}{m_{\rm e} c} \frac{D_\lambda^{\rm max}}{\sqrt{\pi}\xi_{\rm D}}\, q_{jk}(\,T_{\rm ion}) A_k N_N \tag{13.58}$$

and N_N is the column density of nuclei; $D_\lambda^{\rm max}$ the saturation absorption depth as measured in the flux; $\xi_{\rm D}$ is the thermal plus turbulent velocity averaged over the line forming region (see eq. 13.22), and q_{jk} is the mean ionization, averaged over the same region, and described by the temperature $T_{\rm ion}$.

A systematic (practical) method is to build up the curve of growth in several steps. First, one could plot only those lines that originate from the groundlevel of one particular ion. These lines differ in gf-value (see § 7.3) but not in any other parameter. The next step is to plot lines from a different lower level. Because of the excitation term $5039.9\,E_{ljk}/\,T_{\rm ext}$ this new curve of growth will will be shifted relative to the first one. By shifting both curves on top of each other as best as possible the excitation temperature may be derived. The ionization temperature is derived by comparison of the curves for different ions with model predictions. Note that the ionization equilibrium will be a function of temperature and electron density (see eq. 6.24). The latter implies that the shifting due to ionization is, at least in principle, also a measure of the total particle density, and therefore of the surface gravity. Abundances may then be determined by shifting curves of growth for different elements, and comparing these shifts with predictions.

In determining $T_{\rm ext}$, $T_{\rm ion}$, and A_k, using the method described above, relatively weak lines positioned on the linear part of the curve of growth may be used. In this regime the lines are independent of Doppler and micro-turbulent velocity as each photon that can be absorbed by the line will be extincted, regardless of whether it is a line centre or a line wing photon. For saturated lines that are located on the logarithmic (or shoulder) part of the curve of growth this is not so. Saturation implies that (at first) at line centre more potentially extincting particles are present than there are continuum photons offered to the line forming layer. The value of X for which the equivalent width is no longer linear will depend on the value of the micro-turbulent velocity. If $\xi_{\rm micro}$ is relatively large, the flattening of the curve of growth will occur later.

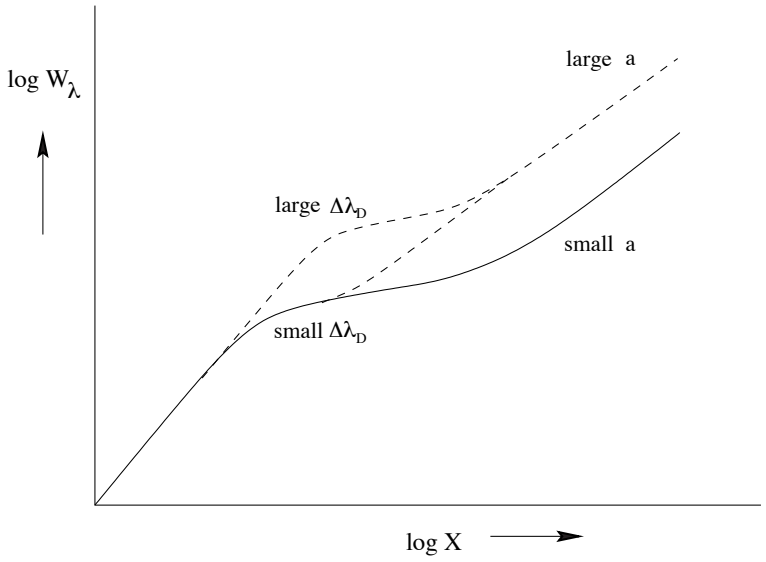


Figure 13.7: *Schematic explanation of the way the curve of growth works. For a large turbulent velocity $\xi_{\rm D}$, proportional to $\Delta\lambda_{\rm D}$, the line will saturated later. Therefore the shoulder part of the curve of growth is further to the right. For a large value of the damping constant, proportional to a, this shoulder part will switch sooner to the $\sqrt{X}$ dependence for strong lines.*

This is so because for large turbulent velocities the intrinsic profile will be broader, therefore less particles can extinct at line centre, such that saturation will occur later, at larger X. This dependence of the curve of growth on $\xi_{\rm micro}$ is pictured in figure 13.7. It may be clear that from a comparison of these regions of the curve of growth with model predictions one may determine the microturbulent velocity.

In principle, one may use the parts of the curve of growth that are caused by the damping wings of strong lines to determine the gravity. In a star with a relatively large value of $\log {\rm g}$ the particle density in the line forming region will be larger. Therefore, the curve of growth of a given line, such as the sodium D_2 in the solar spectrum, will switch sooner from the shoulder part to the $\sqrt{X}$ dependence of strong lines. To use this part of the curve of growth as a diagnostic it is essential that the damping constant γ is well known. After all, γ has a similar effect on the curve of growth as has $\log {\rm g}$, such that a poorly known damping constant will introduce a large uncertainty in the gravity. This is also pictured in figure 13.7.

Exercise 13.1

Show that the equivalent width in frequency units is

$$W_\nu(\text{line}) \equiv \int_{\text{line}} \left(1 - \frac{\mathcal{F}_\nu}{\mathcal{F}_\nu^c}\right) d\nu \tag{13.59}$$

and that it is related to the equivalent width in wavelength units as $W_\nu = c/\lambda^2 W_\lambda$.

Exercise 13.2

Is the equivalent width dependent on interstellar extinction? Why (not)? Is the total line flux dependent on interstellar extinction? Why (not)?

Exercise 13.3

This could be a nice exam question. It is believed that the sun is inside a local bubble of hot gas of about a million degrees and hydrogen density $n_{\text{H}} = 10^{-3}\,\text{cm}^{-3}$. This gas is highly ionized and hence provides no significant opacity for observations within 100 pc or so. However, numerous investigations have shown that there is a minimum column density of neutral hydrogen of about $N_{\text{H}} = 10^{18}\,\text{cm}^{-2}$ toward all stars observed. Therefore it appears that we are embedded in a small, at least partially neutral cloud with approximately this column depth towards its edge. The temperature of the gas in this bubble within a bubble is 8 000 K and the neutral hydrogen density n_{H} = 0.1 cm^{-3}. Assume all H I to be in the ground state. Furthermore, assume that the bubble of neutral gas is spherical and that we are at its center.

a) What is the diameter of this bubble, assuming an atomic hydrogen density of $n_{\text{H}} = 0.1\,\text{cm}^{-3}$.

b) Derive an expression for the optical depth in the Lyα line of hydrogen in terms of the column depth in hydrogen and a constant temperature in the local bubble. Use equation 7.22 and assume a Gaussian line profile function.

c) Is the local bubble optically thin or thick at the center wavelength of Lyα?

d) In spectra of distant quasars we *do* see numerous Lyα lines caused by intervening galactic halos and gas clouds. How can this be?

Exercise 13.4

In this exercise we will derive the Lorentz profile eq. (13.8) following the classical description of Hendrik Antoon Lorentz (1853 – 1928), professor in theoretical physicist at Leiden University and winner of the Nobel prize in 1902. We assume the bound electron in the atom may be described by a (one dimensional) damped harmonic oscillator suffering from periodic excitation with angular frequency ω in an electromagnetic field $E_\circ e^{i\omega t}$. The eigen angular frequency of the damped oscillations is $\omega_\circ$; the damping constant is γ. The equation of motion for this system is therefore given by

$$m_\mathrm{e}\ddot{x} + \gamma m_\mathrm{e}\dot{x} + m_\mathrm{e}\omega_\circ^2 x = eE_\circ e^{i\omega t}, \tag{13.60}$$

where we have introduced $\partial x/\partial t = \dot{x}$ and $\partial^2 x/\partial t^2 = \ddot{x}$ to simplify the notation and e is the charge of the electron. The first term on the left hand side is the force due to inertia; the second term is the damping force, the third term is the restoring force.

a) Using the ansatz $x(t) = x_\circ e^{i\omega t}$, first write down an expression for (the complex solution) $x(t)$ in terms of the unknowns of the problem.

As the solution is complex in nature, Larmor's law (see eq. 8.16) implies that the radiated power is given by

$$p(t) = \frac{2}{3}\frac{e^2}{c^3}\left[\Re(\ddot{x}(t))\right]^2. \tag{13.61}$$

To obtain the mean radiated power $\overline{p}$ requires us to average the radiated power over one period, i.e. to determine $\overline{[\Re(\ddot{x}(t))]^2}$.

b) Derive the expression for $\overline{p}$. Recall that

$$\sin\omega t = \frac{e^{i\omega t} - e^{-i\omega t}}{2i} \quad \text{and} \quad \cos\omega t = \frac{e^{i\omega t} + e^{-i\omega t}}{2} \tag{13.62}$$

You will find that the average power yields an expression of the form

$$\overline{p} = \frac{2}{3}\frac{e^4 E_\circ^2}{3m_\mathrm{e}^2 c^3} f(\nu), \tag{13.63}$$

where $f(\nu)$ is the function describing the frequency dependence of the radiation. To obtain the normalized profile function or Lorentz profile $\phi(\nu)$, fulfilling requirement eq. (7.4), one should determine a normalizing constant C, such that, for instance, $f(\nu) = \phi(\nu)/C$.

c) Determine the constant C and derive the Lorentz profile $\phi(\nu)$, realizing that $\Delta\nu = \nu - \nu_\circ \ll \nu \simeq \nu_\circ$. Recall that

$$\int_{-\infty}^{+\infty} \frac{dx}{1+x^2} = \pi. \tag{13.64}$$

In the 'classical' picture of a driven-damped oscillator, the damping of the system must be exactly compensated by the periodic excitation. This implies that the power associated with the damping force $F_\mathrm{damp} = \gamma m_\mathrm{e}\,\dot{x}(t)$, i.e. $p_\mathrm{damp}(t) = \gamma m_\mathrm{e}\,\dot{x}^2(t)$, and the mean radiated power $p(t) = (2e^2/3c^3)\,\ddot{x}^2(t)$ should cancel out.

d) Use the ansatz $x(t) = x_\circ e^{i\omega t}$ (again) to show that the classical damping constant is given by eq. (13.9).

Exercise 13.5

Before model atmospheres became available it was customary to estimate an electron density from the Inglis-Teller formula which relates the principal quantum number $n_{\mathrm{Balmer}}^{\mathrm{max}}$ of the "last spectrally resolvable line" of the hydrogen Balmer series (i.e. the line with the highest frequency that still can be seen unblended by its neighbouring Balmer lines) to the electron density n_{e} according to the law

$$\log n_{\mathrm{e}} = 23.26 - 7.5 \log n_{\mathrm{Balmer}}^{\mathrm{max}} \tag{13.65}$$

For the A2 I star α Cyg it is found that $n_{\mathrm{Balmer}}^{\mathrm{max}} = 29$; for the A3 V star Sirius one finds $n_{\mathrm{Balmer}}^{\mathrm{max}} = 18$.

a) Explain why there is a relation between n_e and $n_{\mathrm{Balmer}}^{\mathrm{max}}$.

b) Obviously, the electron density is a function of depth in the stellar atmospheres. Which electron density is probed when applying the Inglis-Teller formula: n_{e} at ...

1. the depth at which the continuum optical depth in the optical is about unity.
2. the depth at which the continuum optical depth in the infrared is about unity.
3. the depth at which the optical depth of optical Balmer lines is about unity.
4. the depth at which the optical depth of infrared Balmer lines is about unity.

Explain your answer.

c) Discuss why α Cyg has a higher $n_{\mathrm{Balmer}}^{\mathrm{max}}$ than Sirius.

Exercise 13.6

Derive the rotational broadening profile eq. (13.42). Remember that

$$\int \sqrt{a^2 - x^2}\, dx = \frac{1}{2}\left(x\sqrt{a^2 - x^2} + a^2 \arcsin\frac{x}{a}\right) \tag{13.66}$$

14

Scattering

Arguably the most fundamental *physical* complication inherent in solving radiation transfer problems is that of scattering. Scattering has the effect that the radiation field gets decoupled from local source and sink terms. Because of scatterings, photons may travel large distances through the medium without interacting with this medium. Therefore, due to a strongly deviant radiation field generated at some distant place, the local radiation field may be completely different from the radiation that is produced locally (by the local thermal source function). It is therefore scattering that makes that radiation transfer is a *global* problem and that communicates to regions of large optical depth that the medium has an open edge, through which photons escape. The above states that even at large optical depth, scattering may cause the mean intensity $J_\nu(\mathbf{r}, t)$ to be strongly deviant from the locally produced thermal radiation $B_\nu(T(\mathbf{r}, t))$.

We will investigate the effect of scattering using (1) continuum formation, and (2) line formation in a homogeneous semi-infinite medium. It may be clear that we – in order to determine the effect of scattering – can no longer assume the relevant (continuum or line) source function to be given by the local Planck function. For the line formation example this implies that we may no longer assume LTE. The source function now needs to include a scattering component, therefore it will have the more general form of eq. (9.43). Working out this proto-type source function in some more detail, to get a sense of what to expect, we get

$$S_\nu = \frac{\kappa_\nu B_\nu + \sigma_\nu J_\nu}{\kappa_\nu + \sigma_\nu} = \epsilon_\nu B_\nu + (1 - \epsilon_\nu) J_\nu \tag{14.1}$$

where

$$\epsilon_\nu = \frac{\kappa_\nu}{\kappa_\nu + \sigma_\nu} \tag{14.2}$$

Here ϵ_ν is the *thermal coupling parameter* or *destruction probability*. In case only thermal emission processes play a role, one has $\epsilon_\nu = 1$, therefore $S_\nu = B_\nu$. If only scatterings occur, such that $\epsilon_\nu = 0$, one finds $S_\nu = J_\nu$.

14.1 Continuum scattering

For several types of stellar atmospheres the thermal coupling parameter ϵ_ν for continuum radiation may be very small in a large fraction of the atmosphere. For instance, scattering of light by free electrons is the dominant source of extinction in the outer layers of hot stars. For hot stars even relatively deep layers may show ϵ_ν values of the order of only 10^{-4}, until, finally, for ever increasing density, free-free processes (recall $\chi_\nu^{\rm ff} \propto \rho^2$) win from Thomson scattering (recall: $\chi_{\rm e} \propto \rho$). In cool stars that have a low metal abundance, hydrogen will be neutral in the upper part of the atmosphere and free electrons will be scarce. Here Rayleigh scattering on H and H_2 will dominate the H^- extinction and ϵ_ν will be almost zero up to large depth (when finally hydrogen will rather abruptly excite and ionize, and ϵ_ν will quickly approach unity).

We will now show that if $\epsilon_\nu \rightarrow 0$ the continuum radiation field may (will!) differ from the local Planck function up to large optical depth. To do so we adopt a geometry of planar layers and assume that the Planck function is a linear function of optical depth, i.e.

$$B_\nu(\tau_\nu) = a_\nu + b_\nu \tau_\nu \tag{14.3}$$

Below, we will even assume the medium to be homogeneous, i.e. $b_\nu = 0$. Until then we keep things somewhat more general. We do already assume that ϵ_ν is constant. The $0^{\rm th}$ order moment of the transfer equation (eq. 4.43) can be written as

$$\frac{{\rm d}H_\nu}{{\rm d}\tau_\nu} = J_\nu - S_\nu = \epsilon_\nu (J_\nu - B_\nu) \tag{14.4}$$

If we use the Eddington approximation (we embark on a journey towards large optical depth, where $f_\nu \rightarrow 1/3$) to re-work the $1^{\rm st}$ order moment (eq. 4.44) to

$$\frac{{\rm d}K_\nu}{{\rm d}\tau_\nu} = \frac{1}{3}\frac{{\rm d}J_\nu}{{\rm d}\tau_\nu} = H_\nu \tag{14.5}$$

we find after substitution

$$\frac{1}{3}\frac{{\rm d}^2 J_\nu}{{\rm d}\tau_\nu^2} = \frac{1}{3}\frac{{\rm d}^2 (J_\nu - B_\nu)}{{\rm d}\tau_\nu^2} = \epsilon_\nu (J_\nu - B_\nu) \tag{14.6}$$

where the second equality is allowed because of the particular form of eq. (14.3). The general solution of this differential equation is

$$J_\nu - B_\nu = \alpha_\nu e^{-\sqrt{3\epsilon_\nu}\,\tau_\nu} + \beta_\nu e^{+\sqrt{3\epsilon_\nu}\,\tau_\nu} \tag{14.7}$$

As we require that $J_\nu \rightarrow B_\nu$ for $\tau_\nu \rightarrow \infty$ it must hold that $\beta_\nu = 0$. To determine α_ν we use the boundary condition $H_\nu(0) = g_\nu(0) J_\nu(0)$. The value of the Eddington factor depends on the geometry of the problem. In the case of a sharply peaked radiation field $g_\nu(0) = 1$ (see § 3.6). However, here it is much more appropriate to adopt the solution of the grey

atmosphere: $g_\nu(0) = 1/\sqrt{3}$ (see eq. 10.9 and table 10.1). The boundary condition equation then becomes (see eq. 14.5)

$$J_\nu(0) = \sqrt{3}H_\nu(0) = \frac{1}{\sqrt{3}} \left.\frac{\mathrm{d}J_\nu}{\mathrm{d}\tau_\nu}\right|_0 \tag{14.8}$$

such that we find for the mean intensity

$$J_\nu(\tau_\nu) = a_\nu + b_\nu\tau_\nu \;+\; \frac{b_\nu/\sqrt{3} - a_\nu}{1+\sqrt{\epsilon_\nu}} e^{-\sqrt{3\epsilon_\nu}\,\tau_\nu} \tag{14.9}$$

and for the source function

$$S_\nu(\tau_\nu) = a_\nu + b_\nu\tau_\nu \;+\; (1-\epsilon_\nu)\frac{b_\nu/\sqrt{3} - a_\nu}{1+\sqrt{\epsilon_\nu}} e^{-\sqrt{3\epsilon_\nu}\,\tau_\nu} \tag{14.10}$$

As already implied by eq. (14.1) for small ϵ_ν the source function will be almost equal to the mean intensity. Note that for $\tau_\nu \to \infty$ one recovers $J_\nu \to B_\nu$ and $S_\nu \to B_\nu$, as one should. The above solutions clearly show the physics of the scattering problem. First, at the surface J_ν is remarkably different from B_ν. If we assume for simplicity that the medium is homogeneous, i.e. $b_\nu = 0$, we find that at $\tau_\nu = 0$

$$J_\nu(0) = \frac{\sqrt{\epsilon_\nu}}{1+\sqrt{\epsilon_\nu}} B_\nu \tag{14.11}$$

For very small thermal coupling parameters ϵ_ν one obtains $J_\nu(0) \to \sqrt{\epsilon_\nu}B_\nu$, i.e. the mean intensity becomes much smaller than the Planck function. The run of the source function in a homogeneous medium is given in figure 14.1. It holds that $S_\nu(0) = \sqrt{\epsilon_\nu}B_\nu$ (a result that is much more general than one would expect given the rather rigorous approximations we have made here). For $\epsilon_\nu \ll 1$ we obtain $J_\nu(0) \to S_\nu(0)$. This illustrates that the mean intensity follows the source function, and not the Planck function. We also see that the discrepancy between J_ν and B_ν reached up to large depth in the medium. The exponential term shows that only $J_\nu \to B_\nu$ if $\tau_\nu \gtrsim 1/\sqrt{\epsilon_\nu}$. For media that are dominated by scatterings this can be a very large optical depth. When eventually the mean intensity approaches the Planck function we say that the radiation field is *thermalized*. $\tau_\nu = 1/\sqrt{\epsilon_\nu}$ is called the *thermalization depth*.

The following physical insight may help in gaining a beter understanding of the concept of thermalization optical depth: the thermal coupling parameter $\epsilon_\nu = \kappa_\nu/(\kappa_\nu + \sigma_\nu)$ is simply the probability that a photon that is interacting with a particle is destroyed (i.e. converted into thermal energy). To be absolutely sure that the photon is destroyed it has to experience on the order of $n = 1/\epsilon_\nu$ interactions. If the trajectory of the photon through the medium is described by a *random walk*, with a mean optical photon path $\Delta\tau_\nu \sim 1$ (see § 4.3), the total optical depth it can travel before being destroyed will given by

$$\tau_\nu = \sqrt{n}\,\Delta\tau_\nu = 1/\sqrt{\epsilon_\nu}\,\Delta\tau_\nu \sim 1/\sqrt{\epsilon_\nu} \tag{14.12}$$

Photons that are emitted at a larger optical depth most probably will not be able to escape without begin thermalized first (which is why $J_\nu \to B_\nu$), while photons that are generated

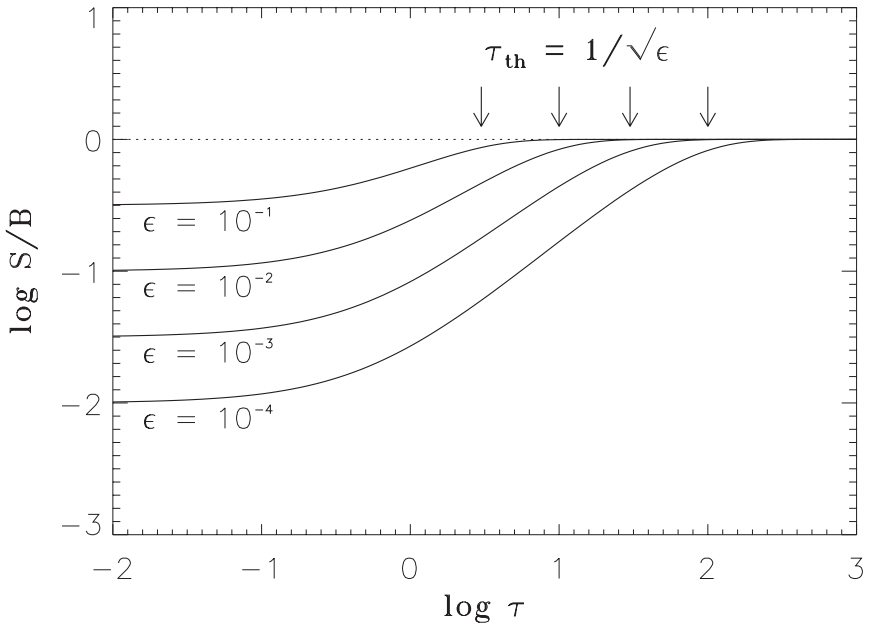


Figure 14.1: *The continuum source function in a homogeneous semi-infinite medium as a function of the optical depth τ_ν. S_ν is normalized to the Planck function B_ν. The curves give solutions for different values of the destruction probability ϵ_ν.*

at a smaller optical depth most likely are able to escape (causing J_ν to become less than the thermal value B_ν). Note that if $\epsilon_\nu = 1$, i.e. in each interaction the photon is destroyed, the thermalization optical depth $\tau_\nu \sim 1$, as expected, because in this case the optical depth scale is the same as the thermal optical depth scale ($d\tau_\nu = d\tau_\nu^{\rm therm} \equiv \kappa_\nu dz$).

14.2 Line scattering in a two-level atom

One of the best examples showing that the source function is composed of a thermal and a scattering contribution is that of the two-level atom. The two-level atom is an idealized problem. Real atoms have many energy levels, therefore, at first sight, this approximation may seem to be grossly inadequate. However it actually provides a surprisingly good description of line formation in a number of cases of interest. This is why we will discuss it here in certain detail.

The general formulation of the line source function in a two-level atom is given by eq. (7.27).

If we assume complete redistribution this reduces to

$$S_\nu^\ell = \frac{n_u A_{ul}}{n_l B_{lu} - n_u B_{ul}} = \frac{2h\nu_{lu}^3}{c^2} \frac{1}{(g_u n_l)/(g_l n_u) - 1} \tag{14.13}$$

The last equality is identical to eq. (7.28). Note that the source function is dependent on the population ratio n_l/n_u. In LTE we assume that the line source function is equal to the local Planck function, i.e. $S_\nu^\ell = B_\nu$, implying that the ratios between level populations are given by the Boltzmann excitation equation. Consequently, it also implies that in LTE scattering processes do not play a role in determining the behaviour of the line source function. By that, the state of the gas can not depend on a radiation field that is generated elsewhere – where conditions may be completely different.

To study the effects of scatterings we must therefore not assume LTE. Rather, we must derive the ratio of the population of lower level 1 and upper level 2 from the statistical equilibrium equation (9.19). This equation states that the number of transitions into state 1 (or 2) should be equal to the number of transitions out of state 1 (or 2); phrased differently

$$n_1 \left(R_{12} + C_{12}\right) = n_2 \left(R_{21} + C_{21}\right) \tag{14.14}$$

Here R is the radiative- and C the collisional rate per particle per second. The radiative excitation rate is given by eq. (7.5); that of radiative de-excitation is the sum of eq. (7.3) and (7.6). The collisional excitation rate is given by eq. (7.7). The relation between C_{12} and C_{21} follows from eq. (7.9), and for our two-level atom is given by

$$\frac{C_{21}}{C_{12}} = \left(\frac{n_1}{n_2}\right)^* = \frac{g_1}{g_2} \exp\left(h\nu_{12}/kT\right) \tag{14.15}$$

Source function of the two-level atom

Armed with the above knowledge, it follows, after substitution of the SE ratio for n_1/n_2 in the outermost right hand side term of equation (14.13), and using the Einstein relations eq. (7.15) and eq. (14.15), that

$$S^\ell = (1-\epsilon)\overline{J}_{12} + \epsilon B_{\nu_{12}} \tag{14.16}$$

where

$$\epsilon = \frac{\epsilon'}{1+\epsilon'} \qquad \text{in which} \qquad \epsilon' = \frac{C_{21}\left[1 - \exp(-h\nu_{12}/kT)\right]}{A_{21}} \tag{14.17}$$

The two-level approximation provides a good approximation for resonance lines. These usually occur in the ultraviolet part of the spectrum where the frequency is relatively high, therefore $h\nu/kT \gg 1$. This implies that ϵ may be approximated by

$$\epsilon \approx \frac{C_{21}}{C_{21} + A_{21}} \tag{14.18}$$

which shows that it simply represents the thermal coupling parameter or destruction probability, i.e. the probability that an absorbed photon is destroyed by a collisional de-excitation process (C_{21}) rather than being conserved by re-emission in a spontaneous de-excitation (A_{21}). Equation (14.16) is the most fundamental equation of the problem. The first term at the right hand side reflects the creation of photons by scattering, i.e. by emission following the absorption of photons. The second term reflects the thermal creation of photons, i.e. by emission following collisional excitations.

We will assume that the profile function is given by a pure Doppler profile, i.e. we neglect natural line broadening and accept that the only broadening is due to the thermal motions of atoms. For a description of this profile we use its dimensionless form eq. (13.20). The absorption in the line can then be described as

$$\chi_x = \chi\phi(x) \tag{14.19}$$

By analogy we obtain for the emission $\eta_x = \eta\phi(x)$. Note that the extinction χ_x is not equal to the extinction at line centre, χ_0. In case of a Doppler profile it holds that $\chi_0 = \chi/\sqrt{\pi}$. The optical depth τ corresponding to the frequency-independent opacity χ is called the *frequency-averaged opacity* in the line, and is the logical optical depth scale in terms of which we will discuss the two-level atom. The optical depth τ and the monochromatic optical depth are related by

$$\tau_x = \tau\phi(x) \tag{14.20}$$

Interpretation of the two-level atom

From a mathematical point of view the source function (eq. 14.16) is still a linear function of the mean intensity. This implies, that although there is a global coupling between the radiation field and the material medium, the problem is still relatively simple and can, for instance, be solved with the Feautrier method. This solution method will be presented in some detail at the end of this section. We first take a closer look at the meaning of the solution of the two-level atom, presented in figure 14.2. To obtain this solution we have used the fact that the medium is homogeneous, ϵ, B_ν, and χ are independent of τ.

What do we learn from this result? In the limit $\tau \to 0$ the source function $S^\ell \to \sqrt{\epsilon}B_{\nu_{12}}$. This result is valid regardless of the exact shape of the profile function. Several rigorous mathematical proofs of this result exist, but these will not be discussed here. For us it is important that $S^\ell(0) < B_{\nu_{12}}$. Why is this so? In a homogeneous medium departures from LTE (i.e. $S^\ell_\nu \neq B_\nu$) arise only because of the presence of a boundary through which photons escape. In regions of high opacity, where the line photons do not yet "feel" the presence of the boundary, all microscopic processes depicted in figure 7.1 are in detailed balance. Phrased differently: LTE holds. However, as soon as photons do start to feel the edge, radiative excitations will no longer be balanced by radiative de-excitations. Since the absorption rate depends on the number of photons present, while the spontaneous emission rate does not (we neglect for simplicity the stimulated emission), the number of radiative excitations drops below the number

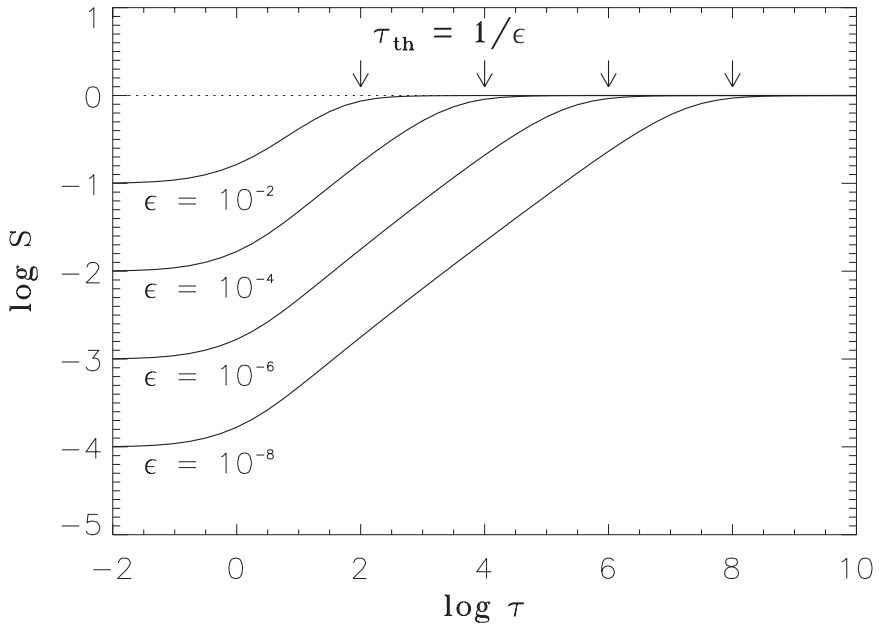


Figure 14.2: *The source function of the two-level atom in a homogeneous semi-infinite medium as a function of the frequency averaged optical depth τ. S is normalized to the Planck function $B_{\nu_{12}}$. The curves give solutions for different values of the destruction probability ϵ, labeled in the figure.*

of de-excitations as soon as photons start to escape. The lower level will consequently start to be depopulated with respect to LTE, while the upper level will be underpopulated. Since the source function measures the number of photons created per unit optical depth, and since the number of created photons is proportional to the population of the upper level (because this is the level from which the atomic transitions accompanied by the photon emission occur), the source function has to drop below the Planck function, i.e. S^ℓ will be less than $B_{\nu_{12}}$.

At some optical depth the source function will start to deviate from the Planck function. This optical depth is called the *thermalization depth* $\tau_{\rm th}$. Figure 14.2 shows that for a Doppler profile $\tau_{\rm th} \sim 1/\epsilon$. Note that for a typical value $\epsilon = 10^{-6}$ the optical depth at which this occurs is very large: $\tau \sim 10^6$. Why is this so? After all, one would expect that the boundary is first felt by an "average" photon when $\tau \sim 1$.

The reason is that in a spectral line it is not the "average" photon that is responsible for the transfer and escape of photons. To better understand this, let us follow the trajectory of a photon from the point where it is created by a thermal emission. We assume that this happens at large optical depth. It is very likely that the newly created photon has a frequency that is close to that of line center. After all, this is where the profile function $\phi(x)$ has its maximum. For this frequency the monochromatic optical depth $\tau_x = \tau\phi(x)$ is large and

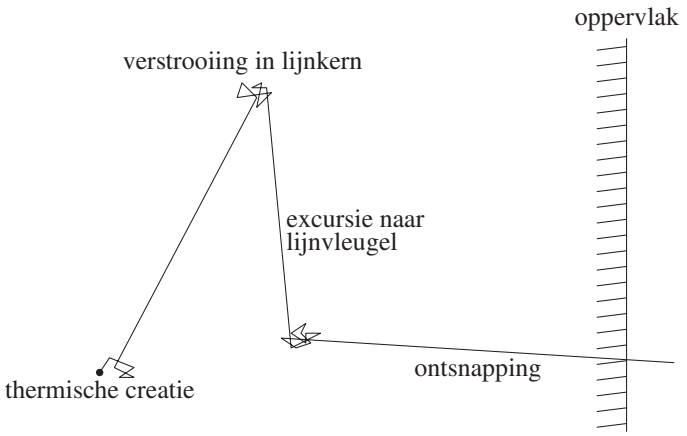


Figure 14.3: *Schematic representation of a trajectory of a photon in a gas of two-level atoms (from Hubeny, in Stellar Atmospheres: Theory and Observation).*

therefore the geometrical distance the photon will travel – i.e. the distance corresponding to $\Delta\tau_x \sim 1$ – will be small (see § 4.3). Probably a similar scenario will unfold in the next scattering events. This yields the following picture of the path a photon will follow in case of complete redistribution: the particle of light will experience many successive scatterings, all at a frequency close to line center; the geometrical distance the photon will cover in all these interactions is, however, small. But, in the rare occasion that the photon is emitted in the wing of the profile function, where the optical depth is orders of magnitude less, it suddenly can travel a large distance. This situation is schematically shown in figure 14.3. Phrased differently: the transfer of photons in the core of the line is very inefficient; the crossing of geometrical space occurs during the rare excursions of the photons into the line wings. It is now clear why the thermalization optical depth is so large: it is determined by photons in the line wings that have an average mean free path that is much larger than that of photons in the line core. However, realize that it is the line core photons that define the frequency mean optical depth τ.

Why is the thermalization depth a function of destruction probability? The total number of successive scatterings is of the order $1/\epsilon$; if the photon has not escaped before it has experienced $1/\epsilon$ scatterings it will be destroyed by means of a collisional deexcitation. It will not have been aware of a boundary to the medium. This illustrates that ϵ must play a role in determining the thermalization depth. But why is $\tau_{\rm th} \propto 1/\epsilon$, and not, for instance, proportional to $1/\sqrt{\epsilon}$ as is to be expected in case the photon distances itself from its point of creation by means of a random walk (see § 14.1)? Essentially, the answer is already given above: if the photon escapes, it does so in the line wing by means of one long flight in the right direction, i.e. in the direction of the surface. (As long as the photon has a frequency close to that of line center it follows a more or less "random walk", however, as discussed, it does not travel a significant geometrical distance.)

We may quantify the above consideration using the *escape probability* formalism.

Escape probability

The probability that a photon with frequency ν and propagating in the direction specified by angle μ escapes in a single flight is given by

$$p_{\nu\mu} = e^{-\tau_{\nu\mu}} \tag{14.21}$$

This follows from the physical meaning of optical depth (see section 4.3). In a plane-parallel medium it holds that $\tau_{\nu\mu} = \tau_\nu/\mu$. The angle-averaged escape probability is thus given by

$$p_\nu = \frac{1}{2}\int_0^1 e^{-\tau_\nu/\mu} d\mu \tag{14.22}$$

where the integration only extends for angles $\mu \geq 0$, since photons moving in the inward direction ($\mu \leq 0$) cannot escape. The angle- and frequency-averaged escape probability for photons in a single spectral line is given by

$$p_{\rm esc} = \int_{-\infty}^{+\infty} \phi_x p_x(\tau_x) dx \tag{14.23}$$

where we have switched to the frequency variable x and the frequency averaged optical depth τ. Notice that at the surface $p_{\rm esc}(0) = 1/2$, because a photon is either emitted in the outward direction, in which case it certainly escapes, or in the inward direction, in which case it does not escape.

We now need to compare the escape probability $p_{\rm esc}$ to the photon destruction probability $p_{\rm des} = \epsilon$. If $p_{\rm esc} \ll p_{\rm des}$, photons are likely thermalized before escaping from the medium. In other words: the line photons de not feel the presence of the boundary, and therefore $S^\ell \sim B_{\nu_{12}}$. If $p_{\rm esc} \gg p_{\rm des}$, photons likely escape before being thermalized. It is therefore natural to define the thermalization optical depth $\tau_{\rm th}$, as

$$p_{\rm esc}(\tau_{\rm th}) \equiv p_{\rm des} = \epsilon \tag{14.24}$$

i.e. as the optical depth for which photon escape (by direct flight) is equally probable as the destruction of photons by a collisional process.

Substitution of (14.23) in (14.24) yields the sought-after relation between $\tau_{\rm th}$ and ϵ. This relation for the escape probability may be a bit too abstract to give a clear physical picture. To obtain a deeper understanding we will approximate the intergral in equation 14.23 using a dichotomous model. To do so we divide the frequencies x in two categories: the "optically thick" frequencies, for which $\tau_x > 1$, and the "optically thin" frequencies, for which $\tau_x < 1$. The *critical frequency* that divides the two regimes is given by

$$\tau\phi(x_{\rm c}) = 1 \tag{14.25}$$

For the Doppler profile (13.20) it follows that

$$x_{\rm c} = \sqrt{\ln(\tau/\sqrt{\pi})} \tag{14.26}$$

i.e. the critical frequency is a function of the frequency averaged optical depth τ. In our dichotomous model we now assume that all photons having an "optically thick" frequency will be absorbed, while all frequencies have an "optically thin" frequency will escape by a direct flight in the outward direction. The photon has a probability $\frac{1}{2}\phi(x)dx$ to be emitted in the outward direction with a frequency in the interval $(x, x+dx)$. In this picture the total escape probability of the photon is given by

$$p(\tau) = \frac{1}{\sqrt{\pi}} \int_{x_{\rm c}}^{\infty} \exp(-x^2)\, dx = \frac{1}{2}{\rm erfc}(x_{\rm c}) \approx \frac{\exp(-x_{\rm c}^2)}{2\sqrt{\pi}x_{\rm c}} \tag{14.27}$$

where erfc(x) is the complimentary error-function. The asymptotic behavior of this function is given by the expression at the far right hand side. Finally, substition of (14.26) in (14.27) yields

$$p(\tau) \approx \frac{1}{2\tau\sqrt{\ln(\tau/\sqrt{\pi})}} \tag{14.28}$$

which within a factor of two (it is twice as large) is equal to the exact asymptotic expression, i.e. if we would not have adopted the dichotomous model. Equating the obtained result to the escape probability ϵ we get

$$\tau_{\rm th} = \frac{1}{2\epsilon\sqrt{\ln(\tau_{\rm th}/\sqrt{\pi})}} \approx \frac{\rm constant}{\epsilon} \tag{14.29}$$

where the constant is of the order unity. This, finally, yields the anticipated result: the thermalization optical depth scales as $\tau_{\rm th} \propto 1/\epsilon$.

Numerical solution using Feautrier

We will now describe a way to numerically solve the two-level atom. In doing so we will assume the line to be formed in a homogeneous semi-infinite medium, i.e. ϵ, B_ν en $\phi(x)$ do not depend on τ. However, each of these parameters may be choosen to be depth-dependent without causing additional complications. The source function (14.16) can be written as

$$S^\ell(\tau) = (1-\epsilon)\int_{-\infty}^{\infty} \phi(x) J(\tau, x) dx + \epsilon B \tag{14.30}$$

$$= (1-\epsilon)\int_{-\infty}^{\infty} \phi(x) dx \int_0^1 u(\tau, \mu, x) d\mu + \epsilon B \tag{14.31}$$

with u the symmetric average of the specific intensity (see equation 5.13). Notice that we have dropped the frequency subscript of the Planck function; B_x will not change significantly over the width of the line profile and may be computed once at, say, $x = 0$.

In order to switch to the difference representation, we introduce a set of optical depth points $\{\tau_d\}, d = 0, ..., D$ for which $\tau_0 < \tau_1 < ... < \tau_D$; a set of angle points $\{\mu_m\}, m = 0, ..., M$; and a set of frequency points $\{\nu_n\}, n = 0, ..., N$. Using this grid notation we obtain for the difference form of the source function

$$S_d = (1 - \epsilon) \sum_{n=0}^{N} a_n \sum_{m=0}^{M} b_m u_{dmn} + \epsilon B \tag{14.32}$$

where a_n and b_m are the quadrature weights of frequency and angle, respectively. In other words

$$\int_{-\infty}^{+\infty} \phi(x) f(x) dx \quad \rightarrow \quad \sum_{n=0}^{N} a_n f(x_n) \tag{14.33}$$

$$\int_{0}^{+1} f(\mu) d\mu \quad \rightarrow \quad \sum_{m=0}^{M} b_m f(\mu_m) \tag{14.34}$$

The transfer equation can again be reduced to the standard 2^{nd} order differential form

$$\mu_m^2 \frac{\mathrm{d}^2 u_{mn}}{\mathrm{d}\tau^2} = \phi_n^2 \left(u_{mn} - S \right) \tag{14.35}$$

The source function connects *all* frequencies and angles. Important to realize is that the source function is a linear function of the specific intensity. The set of equations (14.35) therefore represents a system of linear differential equations.

For ease of notation we can group all angles and frequencies in a single set of grid points i for which $(\mu_i, \nu_i) = (\mu_m, \nu_n)$ for $i = m + n \times M$. We choose to introduce a column vector $\mathbf{u}_d$, with dimension $I = M \times N$, consisting of the angle-frequency components at depth d, i.e. $(\mathbf{u}_d)_i = u_{di}$. Transfer equation (14.35) can then be written as a matrix equation

$$-\mathbf{A}_d \mathbf{u}_{d-1} + \mathbf{B}_d \mathbf{u}_d - \mathbf{C}_d \mathbf{u}_{d+1} = \mathbf{L}_d \tag{14.36}$$

This equation is identical to (5.29), however the meaning of $\mathbf{u}_d$ and the coefficients $\mathbf{A}_d$, $\mathbf{B}_d$ and $\mathbf{C}_d$ is different. In the Feautrier scheme eq. (5.30) $\mathbf{A}_d$ are $\mathbf{C}_d$ are now $(I \times I)$ matrices of which the diagonals contain terms given by the difference form of the transfer equation. $\mathbf{B}_d$ is a full $(I \times I)$ matrix of which the diagonal again contains terms given by difference form of the transfer equation, but which also has off-diagonal terms that represent the integral over the angle-frequency points of the source function (14.32). $\mathbf{L}_d$ is a vector that contains the thermal sources. The total system has a *blok tri-diagonal structure* for which the solution proceeds with a forward-backward recursive sweep, as discussed in § 5.2.

However, to obtain the matrix coefficients $\mathbf{D}_d$ and $\mathbf{E}_d$ (the matrix analogs of (5.32) en (5.33) respectively) requires the actual inversion of the full matrices $\mathbf{F}_d \equiv \mathbf{B}_d - \mathbf{A}_d \mathbf{D}_{d-1}$. In real problems the number of frequency- and angle points I tends to become quite large quite easily making matrix inversions very CPU intensive. Therefore, this method – though very elegant – is often not very practical, not even for linear problems.

The alternative is to solve the problem by means of iteration. In an iterative procedure the approach is to simply solve transfer equation (14.35) by means of the standard Feautrier algorithm, i.e. with A_d, B_d and C_d being scalars for every angle-frequency point seperately, using a *given* source function. In other words: we do a formal solution. The given source function is based on the specific intensity values obtained in the previous iteration. This method is known as *lambda iteration*. The drawback of the lambda iteration method is that in problems in which scattering processes play an important role – as is the case for the two-level atom – the convergence properties of the iteration procedure are very poor. However, there is a method, refered to as *approximate lambda iteration* – see e.g. Chapter 11.2 – that overcomes this problem in normal lambda iteration.

Exercise 14.1

Let us consider a geometry of planar layers. The source function is given by eq. (14.1) and the behaviour of the Planck function by eq. (14.3).

a) Show that the Eddington flux is given by

$$H_\nu = \frac{1}{3} b_\nu - \sqrt{\frac{\epsilon_\nu}{3}} \frac{b_\nu/\sqrt{3} - a_\nu}{1 + \sqrt{\epsilon_\nu}} e^{-\sqrt{3\epsilon_\nu}\,\tau_\nu} \tag{14.37}$$

b) Show that the flux is given by the diffusion approximation (vgl. 4.65) if $\tau_\nu \to \infty$

Exercise 14.2

Let us consider a geometry of planar layers. The source function is given by eq. (14.1) and the behaviour of the Planck function by eq. (14.3). We assume that the medium is homogeneous and that $\epsilon_\nu = 1$.

a) Show that the mean intensity is given by

$$J_\nu(\tau_\nu) = B_\nu \left[1 - \frac{1}{2} e^{-\sqrt{3}\tau_\nu} \right] \tag{14.38}$$

b) Explain why $J_\nu(0) = 1/2\, B_\nu$.

Exercise 14.3

a) Derive the line source function of the two-level atom, i.e. equation (14.16)

15

NLTE mechanisms and models

15.1 LTE versus NLTE

As discussed in § 6.7 one assumes that in LTE the state of the material medium in an elementary volume $dV(\mathbf{r})$ is controlled by the local values of the temperature and density[1]: $n_{ijk} = n_{ijk}(N(\mathbf{r}), T(\mathbf{r}))$ of $n_{ijk}(\rho(\mathbf{r}), T(\mathbf{r}))$ of $n_{ijk}(n_{\rm e}(\mathbf{r}), T(\mathbf{r}))$. The remarkable thing here is that in LTE the radiation field $J(\mathbf{r})$ does not seem to play a role in setting the state n_{ijk} of the gas. In section 6.7 we pointed out that this is because in LTE it is assumed that collisional processes dominate the transitions of electrons between energy levels causing the line source function to equal the local Planck function, i.e. $S_\nu^\ell = B_\nu(T(\mathbf{r}))$. In that case, the general equation for determining n_{ijk}, the statistical equilibrium equation, reduces to the Boltzmann equation for the description of the excitation state (eq. 6.21) and the Saha equation for the description of the ionization state (eq. 6.24). We therefore already arrived at the conclusion that deviations from LTE will arise if

i) collisional processes no longer dominate over radiative processes, *and*

ii) the radiation field deviates from the local Planck function, i.e. $J_\nu(\mathbf{r}) \neq B_\nu(T(\mathbf{r}))$.

The first situation occurs if the density in the medium is low, as the rate of collisions scales with the electron density: $C \propto n_{\rm e}$ (see eq. 7.7). The interstellar medium ($n_{\rm e} \sim 10^{-2} - 10^{-1}\,{\rm cm}^{-3}$) in general and nebulae ($n_{\rm e} \sim 10^3 - 10^6\,{\rm cm}^{-3}$; see also section 2.5) in particular are environments in which the (electron) density is extremely low. Also in chromospheres and coronae and in stellar flares ($n_{\rm e} \sim 10^8 - 10^{10}\,{\rm cm}^{-1}$) around cool stars and in stellar winds the density of the outflowing gas is low enough to cause departures from LTE. The typical nature of these departures are that the absence of an effective (collisional) coupling of the

[1]Note that in LTE there can be (and will be) a non-local *indirect* coupling between the radiation field and the material medium through the equation of radiative equilibrium (see eq. 9.42) and to a lesser extent the equation of hydrostatic equilibrium, which contains a contribution from forces related to the gradient in the radiation pressure (see eq. 11.4).

levels leads to a depopulation of excited levels through spontaneous radiative de-excitations[2] and an overpopulation of the ground level. This raining down of electrons is referred to as *electron cascade*.

Can departures from LTE occur *in* the photosphere of stars? To evaluate this question we use the handy (but rough) relation describing the density at $\tau \sim 1$ in the photosphere that we derived in section 10.2 (see eq. 10.18), i.e.

$$\rho(\tau = 1) \sim \frac{1}{\chi' H} \propto \frac{g}{T} \tag{15.1}$$

Consultation of table B.9 shows that densities are relatively high for cool main-sequence stars, such as the sun ($n_{\rm e} \sim 10^{12} - 10^{15}\,{\rm cm}^{-3}$). This is one of the reasons why for solar-like stars LTE is in general a fine approximation (for a second reason, see below). Also in white dwarfs, which have extreme values for the surface gravity g, the density is such that LTE is a good approximation. Yet even in these cases LTE may break down in the very outer layers of the atmosphere. This situation may not affect the continua and weak lines, which are formed deep in the atmosphere, but may produce deviations in the strengths of strong lines (e.g. hydrogen Balmer and Lyman transitions, the He I triplets, and the Ca H and K lines), mainly in the line cores.

The atmospheres of giants and supergiants (with their low g; yielding $n_{\rm e} \sim 10^{10} - 10^{13}\,{\rm cm}^{-3}$, see again table B.9) require, in general, NLTE treatment of both continuum and line formation. The dividing line is not well defined, but a rule of thumb is that the resonance lines of the dominant ions are likely formed under NLTE conditions for $\log g \lesssim 3.5$ or so for most $T_{\rm eff}$. For the O and early-B dwarfs (with their high T) it is also necessary to adopt NLTE.

Let us now reflect on the second situation. The mean intensity $J_\nu(\mathbf{r})$ will differ from $B_\nu(\mathbf{r})$ if the radiation field at $\mathbf{r}$ is significantly affected by photons that are produced elsewhere, at a place where the conditions are substantially different from those at $\mathbf{r}$. We first consider the situation *in* a photosphere.

In the stellar photosphere

The photospheres of solar-like stars differ fundamentally from those of hot O stars. In solar-like stars the contribution of scatterings to the total extinction is small. In other words, the thermal coupling parameter $\epsilon_\nu = \kappa_\nu/(\kappa_\nu + \epsilon_\nu)$ (eq. 14.1) is almost unity. This is the second reason why LTE works so well in the sun[3]. In O stars scatterings contribute considerably to the total extinction (through Thomson scattering) such that ϵ_ν becomes small. Consequently, by means of scattered photons a radiation field that is produced elsewhere may reach $\mathbf{r}$ and dominate $J_\nu(\mathbf{r})$.

[2] However, levels very close to the continuum remain in LTE. See paragraph COLLISIONAL COUPLING OF HIGH LEVELS WITH THE CONTINUUM in section 8.1.

[3] The two reasons mentioned are so important that the effect of a lower degree of ionization in the solar photosphere ($n_{\rm e}/N_N \sim 10^{-4}$) relative to that in the photosphere of O stars ($n_{\rm e}/N_N \sim 1$) is compensated.

Let us characterize the non-local radiation field by the radiation temperature $T_{\rm R}$. To illustrate the situation we will describe the radiation field in this chapter by assuming that

$$J_\nu(r) = W(r) B_\nu(T_{\rm R}), \tag{15.2}$$

where we have made use of equations (3.11) and (3.12) and of eq. (6.43). $W(r)$ is the geometrical dilution. In the atmospheres of giants and supergiants, and in those of O and early-B dwarfs, it is the condition $J_\nu \neq B_\nu(T)$ that causes departures from LTE. It can be either because $J_\nu > B_\nu(T)$ or $J_\nu < B_\nu(T)$.

Away from the stellar surface

Far away from the stellar surface, in the direction of the interstellar medium, $W(r) \ll 1$. The temperature in the interstellar medium is only $10 - 20$ K in molecular clouds and may reach $\sim 10\,000$ K in gas that is ionized by a nearby O star. This implies that the radiation field $J_\nu(r)$ in the ISM will deviate very strongly from $B_\nu(T_{\rm R})$ and one may expect the state of the gas in the ISM to differ from LTE in important ways.

— — —

In the following sections we provide an overview of the mechanisms of a number of NLTE processes. Though we will deal with the principles of NLTE only, the title of a number of these sections could have been 'NLTE mechanisms as diagnostics of low density media' because a number of these processes *an sich* are interesting as they provide very usable diagnostics for the study of (rarified) gaseous media. In the discussion we especially aim at mechanisms for which only a limited number of transitions need to be considered, such as to provide a clear insight in the NLTE process. We first adress NLTE radiation mechanisms in ionized gas clouds in interstellar space; only then do we consider the situation in stellar atmospheres.

15.2 Photo- and collisional ionization in a low density medium

So one may expect NLTE conditions in a low density medium, such as a gas cloud. Let us consider a gas cloud (often called gaseous nebula) at some distance from a hot star. Such clouds often show emission lines of hydrogen, for instance Hα. The emission arises because the hot star ionizes the nebular gas, after which the free electrons again recombine to arbitrary levels. It is almost certain (see below) that a hydrogen atom in an excited state will de-excite to a lower energy level by means of a spontaneous de-excitation, before it can absorb a photon that may bring it into a higher excited state, or (again) ionizes the atom. As a result, the electrons 'rain down' into the ground level. For instance, an electron that enters the atom at excitation level $i = 6$ may first cascade to level $i = 3$ emitting a Paschen γ photon at a wavelength of 10938 Å (see figure 6.4 en tabel 7.1). Next, it may rain down into level $i = 2$

emitting an Hα λ6562 photon and subsequently end up in the ground level after emitting a Lyα photon.

It may be surprising that a hot star is capable of ionizing nebulae, for the dilution of the radiation fields is considerable even if the cloud is nearby. If we formulate the dilution factor as (see eq. 3.12)

$$W(r) \simeq \frac{1}{4}\left(\frac{R_\star}{r}\right)^2 = 1.272 \times 10^{-16}\left(\frac{R_\star/R_\odot}{d/\mathrm{pc}}\right)^2, \tag{15.3}$$

and the star has a radius of 10 $R_\odot$ and is positioned at a distance of 0.1 pc from a nebula, we find that the geometrical dilution of the radiation field is a factor W = $1.3\,10^{-12}$. So, even though the color temperature of the radiation field may be 30 000 K or higher, the energy density (see eq. 3.15) is weakened by a factor $\sim 10^{12}$. Atomic processes that depend on radiation, such as photoexcitation and photoionization, therefore proceed at a rate that is a factor W lower compared to thermal equilibrium. This explains why excited states are 'almost certain' to experience spontaneous de-excitations rather than photo-excitations. Consequently, the excitation will strongly deviate from the Boltzmann equation (6.21) as the NLTE departure coefficients $b_0 \gg 1$ and $b_i < 1\ \forall_i > 1$ (see vgl. 6.27; but see Sect. 8.1 once $i \sim 100$).

The rate of recombinations is independent of W. Still, this does not mean that the recombination rate is high. The reason is that it is proportional to the electron density n_e. As this is very low 'the gas recombines only laboriously'. This is why in the surroundings of a hot star conditions may arise in which hydrogen is highly ionized. These regions of ionized hydrogen are referred to as H II regions. If the temperature of the star is lower than $\sim$ 30 000 K, the stellar radiation field will simply fail to supply sufficient ultraviolet photons with $\lambda < 912$ Å to ionize the hydrogen gas at all.

Photoionization equilibrium

The above implies that the ionization equilibrium of hydrogen is given by

$$n_0\, 4\pi \int_{\nu_\circ}^{\infty} \alpha_\nu^{\mathrm{bf}} \frac{J_\nu}{h\nu}\, d\nu = N^+ n_\mathrm{e}\, \alpha_\mathrm{A}(T), \tag{15.4}$$

where

$$\alpha_\mathrm{A}(T) = \sum_{i=1}^{\infty} \alpha_{i,\mathrm{H}}^{\mathrm{RR}}(T) \tag{15.5}$$

is the total recombination coefficient and $\alpha_{i,\mathrm{H}}^{RR}$ the recombination coefficient in cm^3 per sec into level i of neutral hydrogen (see eq. 8.8). Values for $\alpha_\mathrm{A}(T)$ are given by, for instance, Osterbrock & Ferland (2006); see Table 18.1. The subscript A indicates Case A recombination; we will return to this in chapter 18.

The ionization fraction $q = n_0/N^+$ (see eq. 6.33). Equation 15.4 describes the photoionization equilibrium, and not the Saha ionization equation 6.24 that specifies this equilib-

rium in LTE conditions. Note that only photo-ionizations, spontaneous recombinations and spontaneous de-excitations feature in the photo-ionization equilibrium.

Collisional ionization equilibrium

Do collisional ionizations play a role in nebulae? Equation (8.13) shows that the rate of collisional ionizations per cm^3 sec^{-1} is proportional to

$$C_{i\kappa} \propto n_{\rm e}\, x^{-2} T^{-3/2} e^{-x}. \tag{15.6}$$

Using eq. (13.56) it follows that

$$x = \frac{\Delta E}{kT} = \frac{E_{\imath\kappa}}{kT} = \frac{11604.8\, E_{i\kappa}({\rm eV})}{T} \tag{15.7}$$

Notice that, as discussed at the end of section 8.1, this rate is a sensitive function of x. For $x \ll 1$ we find that $C_{i\kappa} \propto T^{1/2}$. For $x \gg 1$ the probability of occurrence of collisional ionizations is strongly reduced because of the e^{-x} dependence, reflecting that in this case only a small part of the velocity distribution of the free electrons (given by Maxwell) can be used for collisional ionization processes. For the hydrogen Lyman continuum the ionization energy is 13.6 eV (see figure 6.4) and one finds $x = 157\,825/T$. Applying the effective temperature for T we find that even for the hottest stars x is considerably above unity. In combination with the low electron density ($C_{i\kappa} \propto n_{\rm e}$) it follows that collisional ionizations are not important.

The situation is different in the (proto-typical case of the) solar corona, where the electron density $n_{\rm e} \sim 10^9$ and the temperature $T > 10^6$ K (see section 16.2). For these conditions $x \ll 1$. The rate $C_{i\kappa}$ is therefore orders of magnitude larger than in nebulae. As will be discussed in chapter 16, the corona already starts some thousands of kilometers above the solar surface. At first sight, one might think that therefore the radiation field of the sun might also play a role in (photo)ionizing the corona. But the effective temperature of the sun is only 5 570 K. The regime $\lambda < 912$ Å is thus located far in the Wien part of the spectrum, where the flux has dropped exponentially (see Fig. 6.2). Therefore, photoionization by the relatively 'cold' solar radiation field has no appreciable effect on the ionization balance of the highly charged coronal gas. The solar corona is hence characterized by a *collisional ionization equilibrium*, for hydrogen given by

$$n_0\, C_{0,{\rm H}}(T) = n_0\, n_{\rm e}\, q_{0,{\rm H}}(T) = N^+ n_{\rm e}\, \alpha_{\rm A}(T). \tag{15.8}$$

15.3 Excitations in a low density medium

In the previous section it was discussed that relatively nearby stars, provided their radiation temperature T_R is sufficiently high, are capable of ionizing gaseous nebulae. Is this nebular material also excited, and if so, by what mechanism? We discuss the excitation of gas in a medium of low density using the oxygen O III ion as an example.

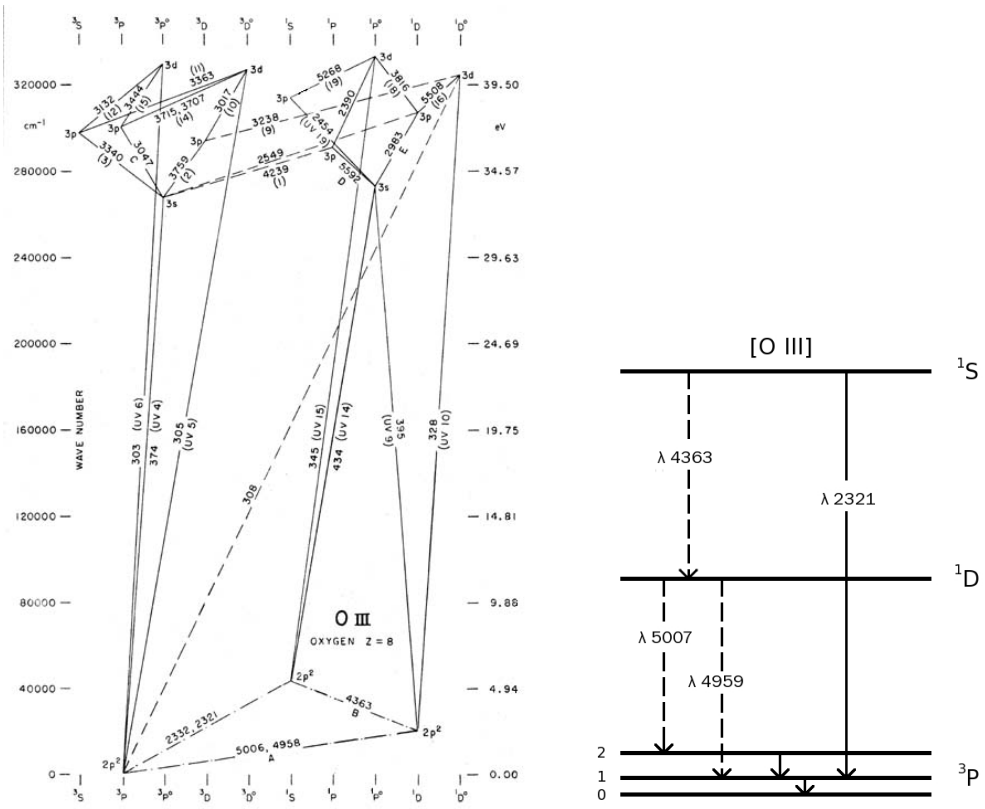


Figure 15.1: *Left: Grotrian or energy level diagram of* O III. *Right: detail of the lowest levels of* O III *and the forbidden oxygen transitions (denoted by dashed lines) that are prominently visible in the emission spectra of nebulae. The* ^{3}P *term of the* $2p^2$ *level has three fine structure levels, i.e.* $^3\mathrm{P}_{0,1,2}$. *The relative energy differences between these fine structure levels have been strongly exaggerated for clarification purposes.*

Forbidden transitions

Figure 15.1 shows the Grotrian diagram or energy level diagram of O III. Note that this ion has two energy levels that are fairly close to the $2p^2\,{}^3$P ground level, i.e. the $2p^2\,{}^1$S level with an excitation potential of 5.25 eV and the $2p^2\,{}^1$D level with excitation potential 2.51 eV. Transitions between the same configuration (here $2p^2$) are called 'forbidden' as they violate *Laporte's parity rule* and are denoted with square brackets, e.g. [O III] λ5007. Forbidden transitions mostly have small Einstein coefficients A_{ul}. For instance, for [O III] λ4959 and λ5007 these are 0.007 and 0.014 sec^{-1}. This implies that the mean lifetime of the $^1\mathrm{D}_2$ level of O III is $< t >= 1/0.021 = 36$ sec (see eq. 7.31). This is extremely long, and these levels are therefore referred to as *meta-stable*. Situations comparable to O III exist for other abundant ions, some of which have even lower transition probabilities.

We consider an H II region with a kinetic (electron) temperature of $T \sim 10\,000$ K and an

electron density $n_{\rm e} \sim 10^3 - 10^6$ cm^{-1}. To describe the rate of collisional excitations we apply Van Regenmorter's formula eq. (7.8), which we repeat for convenience:

$$C_{lu} \propto n_{\rm e}\, x^{-1} T^{-3/2} e^{-x}. \tag{15.9}$$

Here $x = h\nu_{lu}/kT = 11604.8\, E_{lu}({\rm eV})/T$. For the two excited O III levels we find $x = 6.21$ and $x = 2.91$. The Maxwell velocity distribution eq. (6.16) shows that a considerable fraction of the electrons will have energies of a few eV, and even quite some having 5 to 7 eV, but that energies of, for instance, 20 eV will be rare. In other words, there will be a reservoir of free electrons capable of exciting ions to meta-stable levels. The allowed transitions can not profit from this collisional excitation mechanisms as their excitation potentials are high (typically 10 to 50 eV). The collisional ionization of O III requires electrons with energies of at least 55 eV and is excluded as well.

What about radiative excitations in the $2p^2\,{}^3{\rm P} - 2p^2\,{}^1{\rm S}$ and $2p^2\,{}^3{\rm P} - 2p^2\,{}^1{\rm D}$ transitions of O III? There are two reasons why these are unimportant relative to collisional excitations. First, because the interstellar radiation field at these wavelengths is extremely weak. Analogues to eq. (15.2) we may write (see also eq. 7.5)

$$J_\nu(r) = W(r) \int_0^\infty B_\nu(T_R)\phi(\nu)\, d\nu \tag{15.10}$$

where $\phi(\nu)$ is the profile function. Because the profile function is extremely narrow (i.e. it has the width of a spectral line) much less photons will be available for excitation compared to, for instance, that in the Lyman continuum discussed in the previous section. Second, we are here concerned with forbidden transitions. This implies that the chance that a photon from the already very weak radiation field is actually absorbed is very small.

Critical electron density

So, how do collisional processes impact the ^{3}P, ^{1}D, and ^{1}S levels of O III? To assess this, we consider the statistical equilibrium equation (see eq. 9.19) for a two-level system that is not exposed to radiation. We find

$$\frac{n_u}{n_l} = \frac{C_{lu}}{C_{ul} + A_{ul}} = \frac{n_{\rm e}\, q_{lu}}{n_{\rm e}\, q_{ul} + A_{ul}} = \frac{n_{\rm e}\, q_{lu}}{A_{ul}} \left(\frac{1}{1 + (n_{\rm e}\, q_{ul}/A_{ul})} \right) \tag{15.11}$$

Two limiting cases can be considered, with a high and low density of electrons. In the low density limit we find

$$n_{\rm e}\, q_{ul} \ll A_{ul} \quad \rightarrow \quad \frac{n_u}{n_l} = \frac{n_{\rm e}\, q_{lu}}{A_{ul}} \tag{15.12}$$

Similarly, for the high density limit we get

$$n_{\rm e}\, q_{ul} \gg A_{ul} \quad \rightarrow \quad \frac{n_u}{n_l} = \frac{q_{lu}}{q_{ul}} = \frac{g_u}{g_l}\, \exp\left[-h\nu/kT\right] \tag{15.13}$$

The last equality employs for instance eq. 7.10 and links q_{lu}/q_{ul} to the Boltzmann distribution. Therefore the distribution only depends on the (local) temperature and hence will be in (local) TE.

We can now introduce an important concept in nebular gas analysis, namely that of the *critical density*. This is the density for which spontaneous de-excitations are equally important as collisional de-excitations, i.e.

$$n^{\rm crit} \equiv \frac{A_{ul}}{q_{ul}}. \tag{15.14}$$

The physical meaning of the critical density is clear: for $n_{\rm e} < n_{\rm cr}$ collisional de-excitations will be unimportant, and most de-excitations will be spontaneous, resulting in the emission of a photon. The strength of such emission lines will therefore be a good measure of density. For densities much above $n_{\rm cr}$ de-excitation will be mainly through collisions, producing no photons; we expect to approach LTE. The critical density is determined by the quantummechanical properties of the transition under consideration: small Einstein A_{ul} values result in low critical densities. This is why forbidden lines are so important for nebular studies. We can generalize the expression for the critical density to multilevel systems where the critical density now compares radiative transitions to all lower levels with the collisional transitions to all levels

$$n^{\rm crit} = \sum_{l<u} A_{ul} \Bigg/ \sum_{l<u} q_{ul} \ , \tag{15.15}$$

where we have summed over all levels lower than u. The principle is the same: LTE ensues when the density is larger than the critical density.

For the $^1{\rm D}_2$ level of O III the critical electron density $n_{\rm e}^{\rm crit} = 6.8 \times 10^5\,{\rm cm}^{-3}$. It is quite typical in gaseous nebulae that the electron density is below the critical limit, explaining the presence of the (strong) forbidden lines [O III] $\lambda5007$ $\lambda4959$, and $\lambda5007$. The forbidden lines (one could also refer to them as collisionally excited lines) are discussed in more detail in chapter 18. It turns out that they are excellent diagnostics of the electron temperature T and electron density $n_{\rm e}$ in nebular gas.

15.4 Fluorescence

Insofar the discussion of NLTE processes in low density media dealt with excitations, these were collisional excitations. In case the nebular gas is ionized by the ultraviolet radiation field of a nearby star, relatively many photons of the resonance lines of hydrogen and helium will be present. As discussed in section 15.2, such photons will be created in recombination processes of H and He. If the density is not extremely low the nebular medium will be optically thick for H and He Lyman series photons. These photons are 'trapped' and can only escape if they happen to reach the edge of the nebula. The radiation field in these optically thick lines will

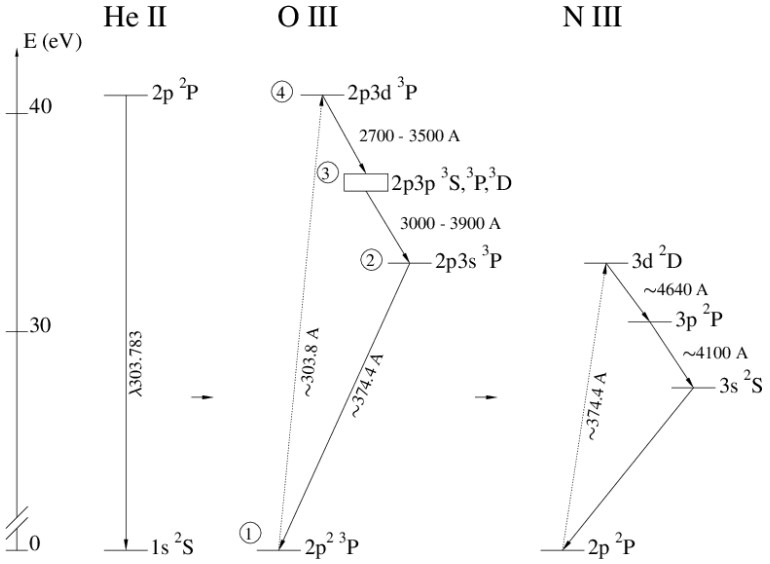


Figure 15.2: *The Bowen fluorescence mechanism: 'resonant' excitation of* O III *lines by He* II.

be much stronger than suggested by equation 15.10, where it is assumed that the medium is optically thin at the line frequencies.

For a nebula that is irradiated by an early O star there will be a zone in the medium in which helium is predominantly present as He III. Obviously, there will be some He II in this zone, after all, there is a photo-ionization balance (*á la* that of hydrogen, see eq. 15.4). In case this balance implies that the amount of He II is such that the He II Lyα photons that are created by the recombination process are scattered many times before they can escape (i.e. the He II Lyα line is optically thick), there will be a strong radiation field at 303.78 Å, the wavelength of the He II Lyα line.

Now again consult the Grotrian diagram for O III in figure 15.1. Notice that, by coincidence, the wavelength of the $2p^2\,{}^3\mathrm{P} - 2p3d\,{}^3P^\circ$ transition is almost identical to that of He II Lyα. To be precise, the most probable transition (having a relative probability of occurence of 0.74) is the $2p^2\,{}^3\mathrm{P}_2 - 3\mathrm{d}\,{}^3\mathrm{P}^\circ_2$ line at 303.80 Å; the next most likely transition (with a relative probability of 0.24) is the $2p^2\,{}^3\mathrm{P}_1 - 3\mathrm{d}\,{}^3\mathrm{P}^\circ_2$ line at 303.62 Å. The many photons at these wavelengths created through the process of recombination and cascading of He III are used to populate the $2p3d\,{}^3\mathrm{P}^\circ$ level from the ground level. Using an alternative phrasing, these transitions are 'resonating' with the He II Lyα line. Many of the de-excitations in O III will proceed through the same transitions back to the ground level (again 'the resonance'), but some will cascade via other (intermediate) levels. This is depicted in figure 15.2. These optical transitions are observed in many planetary nebulae.

With a very small relative probability of 0.02 the $2p3d\,{}^3\mathrm{P}_2$ level (labeled 4 in figure 15.2) will rain down in one of six levels $2p3p\,({}^3\mathrm{S}_1,{}^3\,\mathrm{P}^\circ_{1,2},{}^3\,\mathrm{D}^\circ_{1,2,3})$ (label 3) producing photons with

wavelengths between 2809 Å en 3444 Å. From these levels 14 possibilities arise for further cascading. A number of these occur by way of the $2p3s$ level (label 2), with associated line transitions between 3024 Å en 3811 Å, into the $2p^2\,{}^3\mathrm{P}$ ground level (label 1). This mechanism, producing near-ultraviolet and violet lines between 2809 Å en 3811 Å, is called the *Bowen resonance fluorescence mechanism*, first described by Ira Sprague Bowen (1898 –1973) in 1947.

The line-fluorescence mechanism shows that NLTE effects can be subtle.

Continuum fluorescence

When considering fluorescence one usually means *line* fluorescence. However, it is possible to have a fluorescence effect in rarefied media by continuum radiation from a nearby star. Unlike the case of line fluorescence, in which one specific level is excited, continuum fluorescence is likely to excite a range of levels of an ion (or of multiple ions). The pumping of the levels usually occurs from the ground level through strong allowed transitions. For it to be considered a fluorescence process the radiation field in the line (or lines), $B_{lu}\bar{J}_{lu}$, must dominate over collisional excitations, $n_{\mathrm{e}}q_{lu}$.

An example of continuum fluorescence is the excitation of [Ni II] $\lambda7379$ in the circumstellar ejecta surrounding the Be star P Cygni. The ionization potentials of Fe and Ni are very similar, 7.9 and 16.2 eV for the ground level of Fe I and Fe II, and 7.6 and 18.2 eV for Ni I and Ni II. One may thus expect these ions to co-exist in the same region. It is therefore quite remarkable that the emission in the [Ni II] $\lambda7379$ line in the shell surrounding P Cygni is much stronger than that of apparently comparable lines of [Fe II] in the same spectral regime; sometimes even by up to a factor of 1000 stronger over what is expected on the basis of the abundance ratio between nickel and iron ($A_{\mathrm{Fe}}/A_{\mathrm{Ni}} \sim 20$; see table 16.2).

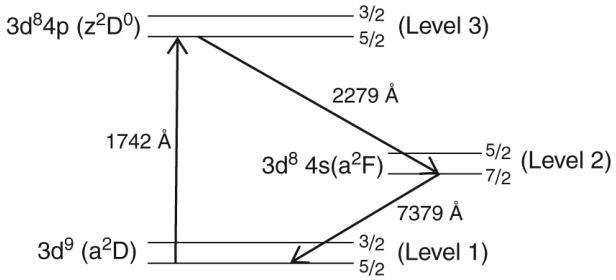


Figure 15.3: *Energy-level diagram illustrating the [Ni II] continuum fluorescence. From: Pradhan & Nahar,* Atomic Astrophysics and Spectroscopy.

The fluorescence mechanism causing the strong [Ni II] $\lambda7379$ Å is schematically drawn in figure 15.3. Multiple channels are possible, but here we focus on the channel shown by the arrows. Continuum photons with a wavelength of 1742 Å from the star P Cygni ($T_{\mathrm{eff}} = 20\,000$ K), at a distance of 0.08 pc from the shell-like nebula, pump electrons from the $a\,{}^2\mathrm{D}_{5/2}$ ground level (level 1 in the figure) to the $z\,{}^2\mathrm{D}^{\circ}_{5/2}$ level (level 3). A spontaneous de-excitation from this level to the excited $a\,{}^2\mathrm{F}_{7/2}$ level (level 2) through the emission of a UV photon of wavelength 2279 Å causes the population in level 2 to increase. The last step in the fluorescence mechanism is the emission of a photon of wavelength 7379 Å in the forbidden transition $a\,{}^2\mathrm{D}_{5/2} - a\,{}^2\mathrm{F}_{7/2}$ when

the electron cascades down from level $2 \rightarrow 1$. Two other forbidden lines in the system drawn in figure 15.3 are pumped in the same way: $a\,^2\mathrm{D}_{3/2} - a\,^2\mathrm{F}_{5/2}$ at 7413 Å and $a\,^2\mathrm{D}_{5/2} - a\,^2\mathrm{F}_{5/2}$ at 6668 Å. The transition $a\,^2\mathrm{D}_{5/2} - a\,^{\mathrm{F}}_{7/2}$ is much weaker.

The critical electron density for this process is $n_{\mathrm{e}}^{\mathrm{crit}} \sim 10^7\,\mathrm{cm}^{-3}$. Only at lower densities continuum fluorescence will occur.

For Fe II the fluorescence process is much less important because the levels corresponding to level 2 in the case of nickel can not be reached by means of allowed transitions (in the case of Ni, transitions $1 \rightarrow 3$ en $3 \rightarrow 2$). Allowed transitions that could correspond to the $1 \rightarrow 3$ transition turn out to have substantially lower critical electron densities compared to the value mentioned for Ni II, which is also unfavorable for the occurrence of fluorescence.

15.5 The NLTE behavior of hydrogen in the atmosphere of O stars

We now study the NLTE behavior of hydrogen *in* the atmosphere of a hot O star. As discussed in the introduction to this chapter we may expect NLTE effects to play a role, especially because of the relatively low gravity. As an example we will focus on a pure-hydrogen planar atmosphere of an O4 V star, with gravity $\log g = 3.6$ (in $\mathrm{cm\,s}^{-2}$) and $T_{\mathrm{eff}} = 45\,000$ K. The run of the NLTE departure coefficients b_i (see eq. 6.27) of the first four levels of hydrogen is given in figure 15.4 as a function of Rosseland optical depth τ_{R} (see section 10.3). Deep in the atmosphere, where densities are high and the contribution of scatterings tot the total source function is small, such that $\epsilon_\nu = \kappa_\nu/(\kappa_\nu + \epsilon_\nu) \rightarrow 1$ (zie vgl. 14.1), the gas is in LTE. For all four levels $b_i = 1$. Moving in the outward direction, i.e. the the direction of decreasing τ_{R}, we first reach the layer in which the continuum is formed. At $\tau_{\mathrm{R}} \sim 0.1$ radiation processes dominate the ionization. To understand the behavior of b_i we consider the ratio of the local radiation field to the local Planck function. Applying approximation eq. (15.2) we may write

$$\frac{J_\nu(r)}{B_\nu(r)} = W(r)\frac{e^{h\nu/kT(r)} - 1}{e^{h\nu/kT_{\mathrm{R}}} - 1} = W(r)\frac{e^x - 1}{e^{x_{\mathrm{R}}} - 1}. \tag{15.16}$$

We may distinguish the following limiting cases

$$\frac{J_\nu(r)}{B_\nu(r)} = \begin{cases} W(r)e^{(x-x_{\mathrm{R}})} & x \gg 1 \\ W(r)(T_{\mathrm{R}}/T) & x \ll 1 \end{cases} \tag{15.17}$$

The value $x = 11604.8\,\Delta E/T$ for the Lyman continuum is $x = 157\,825/T$; for the Balmer continuum $x = 39\,456/T$. We estimate the temperature at $\tau = 0.1$ to be 39 200 K from eq. (10.13). The radiation temperature at this position is about 52 000 K due to scatterings. For the limiting frequency of the Lyman continuum (912 Å) we find, using $x = 4.03$ and $x_{\mathrm{R}} = 3.04$, that $J_\nu/B_\nu = 2.78 \times W(r) > 1$, adopting $W(r) \sim 1/2$. For the limiting frequency of the Balmer continuum (3645 Å) it follows, using $x = 1.01$ and $x_{\mathrm{R}} = 0.76$, that

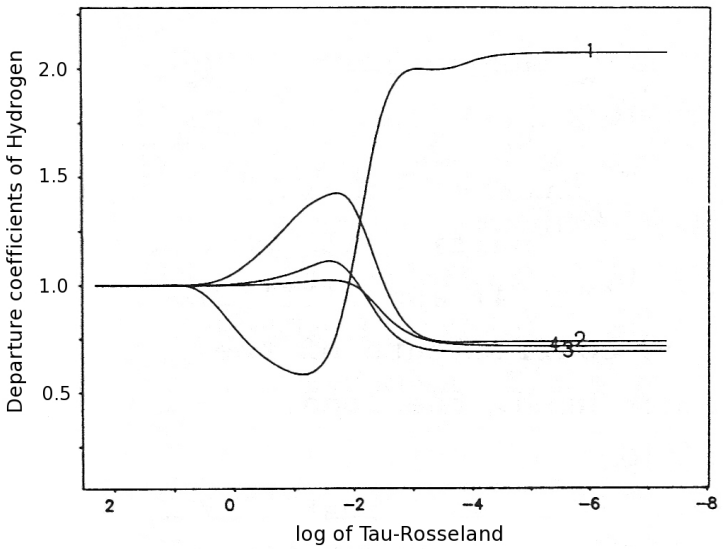


Figure 15.4: *The NLTE departure coefficients of the first four levels of hydrogen as a function of Rosseland optical depth for an O star with effective temperature* $T_{\rm eff} = 45\,000$ *K and gravity* $\log g = 3.6$. *Source: Kudritzki,* The atmospheres of hot stars: modern theory and observation, *1988.*

$J_\nu/B_\nu = 1.53 \times W(r) < 1$, adopting $W(r) \sim 1/2$. The choice $W(r) \sim 1/2$ implies that we consider $\tau_R \sim 0.1$ as 'far from the surface', but that the atmosphere as a whole is compact, i.e. $H/R_\star << 1$ (see section 3.3). For the Paschen continuum ($n_3 \rightarrow \infty$) and the Brackett continuum ($n_4 \rightarrow \infty$) J_ν/B_ν will drop below unity even slightly further. Because the radiation field in the Lyman continuum is stronger than the local Planck radiation field electrons will be pumped from the ground level more efficiently compared to the LTE situation, i.e. the population n_1 will drop below the LTE value, therefore $b_1 < 1$. For the Balmer continuum the situation is reversed. Here the radiation field is weaker compared to LTE, which leads to a (very weak) overpopulation of n_2, i.e. $b_2 > 1$. For n_3 and n_4 this overpopulation is more prominent.

There is a second reason as to why the departure coefficient b_2 does not reach as high a values as that of the two higher levels: the Lyman lines are still optically thick at $\tau_{\rm R} \sim 0.1$ and in particular the levels 1 and 2 are strongly coupled to each other, the oscillator strength f_{12} being large compared to f_{13} and f_{14} (see table 7.1). Because of this the n_1 level has the tendency to impose its departure coefficient on n_2. If this coupling would be complete – one would then state that the Lyα is in *detailed balance* – the value of b_2 would even follow that of b_1. The latter is not the case in figure 15.4, but there is some degree of coupling. This explains why b_2 does not rise as high as b_3, which in turn does not rise as high as b_4.

Proceeding further in the outward direction one will reach the regime where the Lyman lines become optically thin. This happens at $\tau_{\rm R} \sim 0.01$. High levels are now no longer (signif-

icantly) populated by means of photoexcitations from the ground level. The main channel through which these high levels attain electrons is by photoionisation from the ground level followed by recombination and cascading to lower levels. The processes governing these layers therefore are similar to those of the photoionisation equilibrium discussed in section 15.2. The upper levels will rain down in the ground level, causing $b_1 > 1$ en $b_i < 1 \; \forall_i > 1$.

Statistical equilibrium, an example

The simple hydrogen model of the O4 V star lends itself well for a general discussion of the solution of the equations of statistical equilibrium 9.19. To keep a clear view of the situation (and to make sure that the formulae fit on one page) we reduce the number of bound levels of hydrogen to three. For the principle of things this does not matter. Figure 15.5 provides the energy level diagram of our model atom. As always the index κ represents the continuum level, in this case ionized hydrogen. In the general sense n_κ is the ground level of the ion that forms if the ion under consideration is ionized. It therefore could have been the ground level of Fe XX if we were to consider the model for Fe XIX.

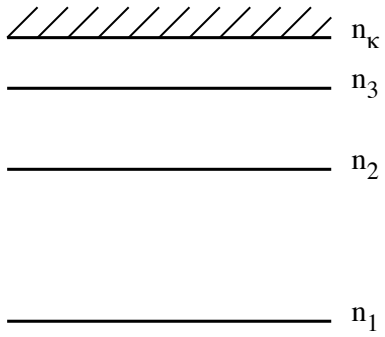


Figure 15.5: *A simple model of an H-atom consisting of three bound levels and a continuum level.*

We assume that the local temperature $T(z)$, the local density $\rho(z)$ and the local radiation field $J_\nu(z)$ are known. As in the LTE model (see sections 11.1 and 11.2) the construction of a NLTE model requires an iterative approach. The word 'known' in the one but last sentence therefore implies: to use the current values of the named parameters. For the hydrogen model the density can be easily coupled to the number density of nuclei through $\rho(z) = m_{\rm H} N_N(z)$.

For our four level atom we may write the following four statistical equilibrium equations:

$$\begin{aligned}
-n_1\left(P_{12}+P_{13}+P_{1\kappa}\right)+n_2P_{21}+n_3P_{31}+n_\kappa P_{\kappa 1} &= 0\\
-n_2\left(P_{21}+P_{23}+P_{2\kappa}\right)+n_1P_{12}+n_3P_{32}+n_\kappa P_{\kappa 2} &= 0\\
-n_3\left(P_{31}+P_{32}+P_{3\kappa}\right)+n_1P_{13}+n_2P_{23}+n_\kappa P_{\kappa 3} &= 0\\
-n_\kappa\left(P_{\kappa 1}+P_{\kappa 2}+P_{\kappa 3}\right)+n_1P_{1\kappa}+n_2P_{2\kappa}+n_3P_{3\kappa} &= 0
\end{aligned} \tag{15.18}$$

where

$$P_{ij}(z) = R_{ij}(z) + C_{ij}(z). \tag{15.19}$$

The coefficients R_{ij} are given by equations (7.3), (7.6), (8.8) en (8.10) for $i > j$ and by (7.5) and (8.1) for $i < j$. The coefficients C_{ij} are given by (7.10) and (8.15) for $i > j$ and by (7.7) and (8.12) for $i < j$.

This system of four equations is however redundant (the determinant equals zero) as one may always reconstruct one of the equations from the remaining three. This is not the only problem, we also still have to fulfill the requirement of particle conservation eq. (6.30). Both problems can be resolved by, for instance, replacing the last equation by

$$n_1 + n_2 + n_3 + n_\kappa = A_k N_N \tag{15.20}$$

where A_k is the number abundance of element k and N_N is the nucleon density of nuclei. The system of equations may be written in matrix format

$$\begin{pmatrix} -P_{12} - P_{13} - P_{1\kappa} & +P_{21} & +P_{31} & +P_{\kappa 1} \\ +P_{12} & -P_{21} - P_{23} - P_{2\kappa} & +P_{32} & +P_{\kappa 2} \\ +P_{13} & +P_{23} & -P_{31} - P_{32} - P_{3\kappa} & +P_{\kappa 3} \\ 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} n_1 \\ n_2 \\ n_3 \\ n_\kappa \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ A_k N_N \end{pmatrix}$$

or

$$\boldsymbol{\mathcal{A}}\,\vec{\boldsymbol{n}} = \boldsymbol{\mathcal{B}} \tag{15.21}$$

where $\boldsymbol{\mathcal{A}}$ is the transition probability matrix or *rate matrix* with dimension 4×4; $\vec{\boldsymbol{n}} = (n_1, n_2, n_3, n_\kappa)$ is the population vector, and $\boldsymbol{\mathcal{B}} = (0, 0, 0, A_k N_N)$ is a vector for which all elements equal zero save one. The last element of this vector provides the number density of element k. In our hydrogen model the number abundance $A_k = A_1 = 1$. The values for the populations follow from

$$\vec{\boldsymbol{n}} = \boldsymbol{\mathcal{A}}^{-1}\,\boldsymbol{\mathcal{B}}. \tag{15.22}$$

Inversion of the matrix $\boldsymbol{\mathcal{A}}$ can for instance be done using a LU decomposition technique.

15.6 NLTE Model atmospheres

An overview of modern NLTE model atmospheres for hot luminous stars is provided in Table 15.1. Several of these codes also account for a radially streaming trans-sonic stellar wind.

Code	DETAIL / SURFACE[1]	TLUSTY[2]	CMFGEN[3]	FASTWIND[4]	PHOENIX[5]	PoWR[6]	WM-basic[7]
Major application	Hot stars with negligible winds	Hot stars with negligible winds	OB(A)-stars WRs, SNe	OB-stars, early A-sgs	Stars cooler than 10 kK, SNe	WRs, O-stars	Hot stars with dense winds, SNe
Geometry	Planar layers	Planer layers	Spherical	Spherical	Spherical	Spherical	Spherical
Diagnostic range	No limitations	No limitations	No limitations	No limitations	No limitations	No limitations	Taylored for UV
Line Radiative Transfer	Observer's frame	Observer's frame	CMF	CMF / Sobolev	CMF / Observer's frame	CMF	Sobolev
Line Blanketing	LTE	Yes	Yes	Approximated	Yes	Yes	Yes
Temperature structure	Radiative equilibrium	Radiative equilibrium	Radiative equilibrium	e^- thermal balance	Radiative equilibrium	Radiative equilibrium	e^- thermal balance
Comments	No wind (Hydrostatic Equilibrium)	No wind (Hydrostatic Equilibrium)	Start model required	User-specified atomic models. Optically thin/thick wind clumping / porosity / vorocity	Molecules included	Optically thin wind clumping / porosity	No wind clumping
Execution time	Few minutes	Hour(s)	Hours	Few minutes to 1h	Hours	Hours	1-2h

Table 15.1: *Original references. For the latest updates on these codes, see elsewhere in these lecture notes and the internet.* [1] *Giddings (1980),* [2] *Hubeny (1998);* [3] *Hillier and Miller (1998);* [4] *Puls et al. (2005);* [5] *Hauschildt (1992);* [6] *Gräfener et al. (2002);* [7] *Pauldrach et al. (2001). Adapted from: Martínez-Núñez et al. (2017).*

16

The sun

16.1 Introduction

The study of the solar atmosphere is fundamental for multiple reasons: *i)* First, the properties of the sun are vital to everything that happens in our solar system. The climate of planetary atmospheres, for instance, is determined by the precise shape of the solar spectral energy distribution. Realize too that our eyes have evolved to be sensitive to the part of the solar spectrum reaching the earth surface in which the most flux is emitted. *ii)* Second, obviously, to learn more about the sun itself. For instance, to learn about its temperature and pressure stratification; chemical composition, magnetic fields, and time dependent spatial structures and velocity fields present at its surface. The basic solar properties are essential for determining its current evolutionary stage. This is done through a confrontation of these properties with models of stellar evolution, which also allows to study the history of our star as well as its future. *iii)* Third, because the sun is by far the brightest star in the sky. It offers the opportunity to study a plethora of processes that also occur in other stars, but at a much higher spatial resolution, a higher signal-to-noise ratio – and related to this a much higher spectral resolution – and a much higher time resolution. *iv)* Fourth, because the solar spectrum is the most stable calibration source of spectral properties of atoms and molecules. Moreover, it provides such information at temperatures that can not be (easily) achieved in laboratory experiments.

One can think of other reasons why the sun is so important – it is for instance the only star for which the equilibrium neutrino luminosity can be measured – however, the above listed reasons should be sufficient.

16.2 General information about the sun

In order to determine the fundamental parameters of the sun it is essential that its distance is known. The mean distance, or *astronomical unit*, is $149\,597\,870 \pm 2$ km. This value is the

mass	$M_\odot$	1.9889 ± 0.0003	$\times 10^{33}$	gr
radius	$R_\odot$	6.957 ± 0.001	$\times 10^{10}$	cm
luminosity	$L_\odot$	3.828 ± 0.010	$\times 10^{33}$	$\mathrm{erg\,sec^{-1}}$
effective temperature	T_{eff}	$5\,772 \pm 4$		K
age		4.57 ± 0.05	$\times 10^{9}$	year
bottom convection zone	r_{cv}	0.713 ± 0.003		$R_\odot$
mass fraction at surface				
– hydrogen	$\mathrm{X_s}$	0.739 ± 0.005		
– helium	$\mathrm{Y_s}$	0.248 ± 0.005		
– other elements	$\mathrm{Z_s}$	0.012 ± 0.002		

Table 16.1: *Basic solar parameters.*

mean distance between the centers of the sun and the earth. The "mean" is over the eccentric anomaly, i.e. the angle between the sun-earth line and the direction to the perihelion of the elliptical orbit of the earth around the sun.

Basic parameters

Fundamental or *basic parameters* of the sun are given in table 16.1. Here, the solar radius is defined as the point in the atmosphere where the visual optical depth in the radial direction $\tau_V = 2/3$. What is actually measured is the angular diameter between the inflection points, at both sides of the solar disk, of the visual specific intensity profile across the solar disk. Model computations show that this is about 300 km above the layer where $\tau_V = 2/3$, for which the radius is corrected. The earth atmosphere absorbs and reflects a sizable fraction of the solar light. For this reason the solar luminosity needs to be measured from space. Balloon, rocket, and satellite observations show both short and long term variations in the bolometric luminosity. Most pronounced is an 11 year cycle in which the maximum of the irradiance corresponds to the number of sunspots. The peak-to-peak amplitude of this variation is about 0.09 percent. The mean luminosity and radius imply an effective temperature of 5 772 K. The mass and radius imply a surface gravity $g = 2.742 \times 10^4$ cm sec^{-2} or $\log g = 4.438$. The age of the sun follows from radioactive decay of long living isotopes, such as U^{238}, in meteorites. The depth of the convection zone is determined from helioseismic measurements.

Abundances

Knowledge of the chemical composition of the sun is essential for the modeling of its atmosphere and internal structure. The sun, though being the best studied star, is considered an ordinary star. This makes it very suited to serve as a reference in abundance studies of other stars. It is also very interesting to compare the chemical composition of other objects

in our solar system – such as planets, comets and meteorites – with those of the sun, as this may give insight in the formation and evolution of the solar system as a whole. Other reasons for studying the solar abundance pattern in great detail is its importance for theories of nucleosynthesis, as well as for models of the chemical evolution of galaxies, our Milky Way in particular, and of the universe itself. The first to perform a "quantitative" analysis of abundances in the solar atmosphere was Henry Norris Russell, who in 1929 on the basis of by-eye estimates determined the abundances of 56 elements.

The signature of 65 of the 83 stable elements is present in the photospheric absorption line spectrum, that is observed at high signal-to-noise ratio from ultraviolet to far infrared wavelengths using high spectral resolution instruments. The remaining elements are also present in the sun, but their lines are too weak to measure in the mentioned spectral range. Surprisingly, helium is in the list of non-observable elements. Spectral analysis of the elemental lines can be done well using LTE model atmospheres because in the solar photosphere collisional processes dominate over radiative processes. The derived abundances are usually expressed in logarithm of the number abundance normalized to 10^{12} particles of hydrogen, i.e.

$$A_{\rm el} = \log(N_{\rm el}/N_{\rm H}) + 12.0 \tag{16.1}$$

Table 16.2 lists these abundances for the most important elements. It is to be noted that the above definition differs from that of A_k in eq. (6.30). The normalization relative to $10^{12}\ N_{\rm H}$ was originally adopted in order to avoid negative numbers for the least abundant elements. Alas, it later turned out that tantalum ($A_{\rm Ta} = -0.13$) and uranium ($A_{\rm U} = -0.52$) are so rare that they do not fulfill this rule.

Nowadays, extremely sophisticated models of the solar atmosphere are used to derive the abundances. In 2005 Asplund and co-workers reported $A_{\rm el}$ values based on a 3-dimensional, time-dependent, hydrodynamical solar model atmosphere (see table 16.2). Compared to earlier work, significantly lower C, N, and O abundances were found. These low C, N, and O abundances agree much better with the corresponding abundances in the local interstellar medium and nearby B stars as well as the values measured in the solar corona/wind. A major challenge for these revised solar abundances, however, is that the excellent agreement achieved until now with helioseismology using standard solar evolution modeling is lost. The solution to this problem has not yet been identified.

It is common to use square brackets to denote that a stellar abundance is given relative to the solar value. The definition is

$$[X] \equiv \log X_{\rm star} - \log X_{\rm Sun} \tag{16.2}$$

In particular this is done for iron, where

$$[{\rm Fe/H}] = \log\left(N_{\rm Fe}/N_{\rm H}\right)_{\rm star} - \log\left(N_{\rm Fe}/N_{\rm H}\right)_{\rm Sun} \tag{16.3}$$

is sometimes referred to as the *metal abundance* of the star. Do not get confused with $Z = 1 - X - Y$, i.e. the mass fraction of all elements except hydrogen (X) and helium (Y), which is

	Element	Symbol	Atomic mass	Photosphere A_k	Mass fraction	Mass Z_s	Meteorites A_k	Phot - Meteor
1	Hydrogen	H	1.0079	12.00	0.735		8.25 ± 0.05	3.75
2	Helium	He	4.0026	10.93 ± 0.01*	0.248		–	–
3	Lithium	Li	6.941	1.05 ± 0.10	5.68(-11)	–	3.25 ± 0.06	-2.20
4	Beryllium	Be	9.0122	1.38 ± 0.09	1.57(-10)	–	1.38 ± 0.08	0.00
5	Boron	B	10.811	2.70 ± 0.20	3.96(-9)	–	2.75 ± 0.04	-0.05
6	Carbon	C	12.011	8.39 ± 0.05	2.15(-3)	126.76	7.40 ± 0.06	+0.99
7	Nitrogen	N	14.007	7.78 ± 0.06	6.15(-4)	36.28	6.25 ± 0.07	+1.53
8	Oxygen	O	15.999	8.66 ± 0.05	5.33(-3)	314.53	8.39 ± 0.02	+0.27
9	Fluorine	F	18.998	4.56 ± 0.30	5.03(-7)	0.03	4.43 ± 0.06	+0.13
10	Neon	Ne	20.180	7.84 ± 0.06	1.02(-3)	60.11	–	–
11	Sodium	Na	22.990	6.17 ± 0.04	2.48(-5)	1.46	6.27 ± 0.03	-0.10
12	Magnesium	Mg	24.305	7.53 ± 0.09	6.01(-4)	35.43	7.53 ± 0.03	0.00
13	Aluminium	Al	26.982	6.37 ± 0.06	4.61(-5)	2.72	6.43 ± 0.02	-0.06
14	Silicon	Si	28.086	7.51 ± 0.04	6.62(-4)	39.06	7.51 ± 0.02	0.00
15	Phosphorus	P	30.974	5.36 ± 0.04	5.17(-6)	0.31	5.40 ± 0.04	-0.04
16	Sulpher	S	32.066	7.14 ± 0.05	3.23(-4)	19.04	7.16 ± 0.04	-0.02
17	Chlorine	Cl	35.453	5.50 ± 0.30	8.17(-6)	0.48	5.23 ± 0.06	-0.03
18	Argon	Ar	39.948	6.18 ± 0.08	4.40(-5)	2.59	–	–
19	Potassium	K	39.098	5.08 ± 0.07	3.43(-6)	0.20	5.06 ± 0.05	+0.02
20	Calcium	Ca	40.078	6.31 ± 0.04	5.96(-5)	3.52	6.29 ± 0.03	+0.02
21	Scandium	Sc	44.956	3.05 ± 0.08	3.68(-8)	–	3.04 ± 0.04	+0.01
22	Titanium	Ti	47.88	4.90 ± 0.06	2.79(-6)	0.17	4.89 ± 0.03	+0.01
23	Vanadium	V	50.942	4.00 ± 0.02	3.71(-7)	0.02	3.97 ± 0.03	+0.03
24	Chromium	Cr	51.996	5.64 ± 0.10	1.59(-5)	0.94	5.63 ± 0.05	+0.01
25	Manganese	Mn	54.938	5.39 ± 0.03	9.83(-6)	0.58	5.47 ± 0.03	-0.08
26	Iron	Fe	55.847	7.45 ± 0.05	1.15(-3)	67.68	7.45 ± 0.03	0.00
27	Cobalt	Co	58.933	4.92 ± 0.08	3.57(-6)	0.21	4.86 ± 0.03	+0.04
28	Nickel	Ni	58.693	6.23 ± 0.04	7.27(-5)	4.29	6.19 ± 0.03	+0.04
29	Copper	Cu	63.546	4.21 ± 0.04	7.51(-7)	0.04	4.23 ± 0.06	-0.02
30	Zinc	Zn	65.39	4.60 ± 0.03	1.90(-6)	0.11	4.61 ± 0.04	-0.01
31	Gallium	Ga	69.723	2.88 ± 0.10	3.86(-8)	–	3.07 ± 0.06	-0.19
32	Germanium	Ge	72.61	3.58 ± 0.05	2.01(-7)	0.01	3.59 ± 0.05	-0.01
33	Arsenic	As	74.922	–	1.28(-8)	–	2.29 ± 0.05	–
34	Selenium	Se	78.96	–	1.48(-7)	0.01	3.33 ± 0.04	–
35	Bromine	Br	79.904	–	2.48(-8)	–	2.56 ± 0.09	–
36	Krypton	Kr	83.80	3.28 ± 0.08	1.16(-7)	0.01	–	–
37	Rubidium	Rb	85.468	2.60 ± 0.15	2.48(-8)	–	2.33 ± 0.06	+0.27
38	Strontium	Sr	87.62	2.92 ± 0.05	5.31(-8)	–	2.88 ± 0.04	+0.04
39	Yttrium	Y	88.906	2.21 ± 0.02	1.05(-8)	–	2.17 ± 0.04	+0.04
40	Zirconium	Zr	91.224	2.59 ± 0.04	2.59(-8)	–	2.57 ± 0.02	-0.02
56	Barium	Ba	137.33	2.17 ± 0.07	1.48(-8)	–	2.16 ± 0.03	+0.01
63	Europium	Eu	151.96	0.52 ± 0.06	3.66(-10)	–	0.49 ± 0.04	+0.03
90	Thorium	Th	232.04	–	2.08(-10)	–	0.06 ± 0.04	–
92	Uranium	U	238.03	$<$-0.47	5.49(-11)	–	-0.52 ± 0.04	–

Table 16.2: *Elemental abundances as determined from the solar photosphere and meteorites given in number ratio relative to hydrogen:* $A_{\rm em} = \log N_{\rm el}/N_{\rm H} + 12$. ** The helium abundance is the current value, and corresponds to* Y_s *= 0.248 (see also table 16.1). For the initial value, i.e. at the time of formation of the sun,* $A_{\rm He} = 10.99 \pm 0.02$, *which corresponds to* Y_s *= 0.275 and* X_s *= 0.708. From: Asplund et al. 2005.*

commonly refered to as the *metallicity*. Relatively young stars in the plane of the Milky Way have about solar abundances (therefore $[\mathrm{Fe/H}] \sim 0$). "Metal poor" stars have $[\mathrm{Fe/H}] \sim -1$ to -2. Population II halo dwarf stars have formed relatively early in the history of the Galaxy and are "extremely metal poor", showing $[\mathrm{Fe/H}] \sim -3$. Astronomers are vigorously searching for stars with even less metals because these must have formed in the very early history of the Milky Way. The abundance pattern of such hyper metal poor stars can tell us about the properties of the first generations of stars in our galaxy, that must have been responsible for the metals in these hyper metal poor objects. In 2002 Norbert Christlieb and collaborators discovered that HE 0107-5240 (Christlieb's star) has an iron abundance [Fe/H] = -5.3 (see figure 16.1). Just as a side note, the mass of Christlieb's star is only 0.8 $M_\odot$. In 2005, Anna Frebel and collaborators discovered He1327-2326 (Frebel's star) to have an even lower iron abundance, [Fe/H] = -5.6. In 2014, the so far lowest metal abundance star was found, SMSS J031300.36-670839.3, by Keller et al. (2014). The star is estimated to have an age of 13.6 Gyr, i.e. to have formed some 0.1–0.2 Gyr after the Big Bang. The gas from which stars that have [Fe/H] $<$ -4.5 are composed are likely enriched by only a single supernovae. The abundance pattern of Keller's star is quite peculiar, with [C/H] = -2.6, [Mg/H] = -3.8, [Ca/H] = -7.0, and [Fe/H] = -7.1. This abundance pattern is best reproduced if prior to the formation of this star a supernova explosion of an initially 60 $M_\odot$ star occured in the immediate surroundings. It is expected to have been a low-energy supernova, one that directly produces a black hole.

Although most of the matter of the solar system is concentrated in the sun, very useful information concerning the chemical composition of the initial solar nebula can be obtained from the study of other bodies of the solar system. A complicating factor in the study of such objects is that they may have suffered chemical fractionation, which is the sinking towards the center of relatively heavy elements at times when such bodies were partly or fully melted. Objects that experienced chemical fractionation include the terrestrial planets and (the parent bodies of) meteorites. For this reason, earth is not suited for doing a representative abundance analysis. The most suited objects turn out to be meteorites of the type *CI carbonaceous chondrites*. These are formed early on in the formation of the solar system and appear to have preserved the bulk composition of the nebula from which they condensed, except for the highly volatile elements (H, C, N, O, and the rare gases) which partly escaped. Meteoritic chemical composition can be measured with such high accuracy that it is better known than the sun's photospheric abundance pattern (of which the abundance of some elements suffer from poorly known oscillator strengths). Meteoritic abundances are also given in table 16.2.

The abundance determination of helium warrants extra attention. In 1868, a new element was discovered in the solar chromospheric spectrum obtained during an eclipse. The name of the sun was given to this unknown element: helium. It was only discovered on earth in 1895. The reason why helium is not present in the photosphere is because in a medium with a typical temperature around 6000 K, no lines of helium (neutral or once-ionized) fall within the visual spectral range. Until recently the He abundance was most commonly derived from chromospheric observations. Now A_{He} is also being determined from stellar structure calculations. This yields $N_{\mathrm{He}}/N_{\mathrm{H}} = 9.8 \pm 0.4$ % (or $A_{\mathrm{He}} = 10.99$) for the initial composition of the sun and the proto-solar nebula. Inversion of helioseismic observations leads to a more

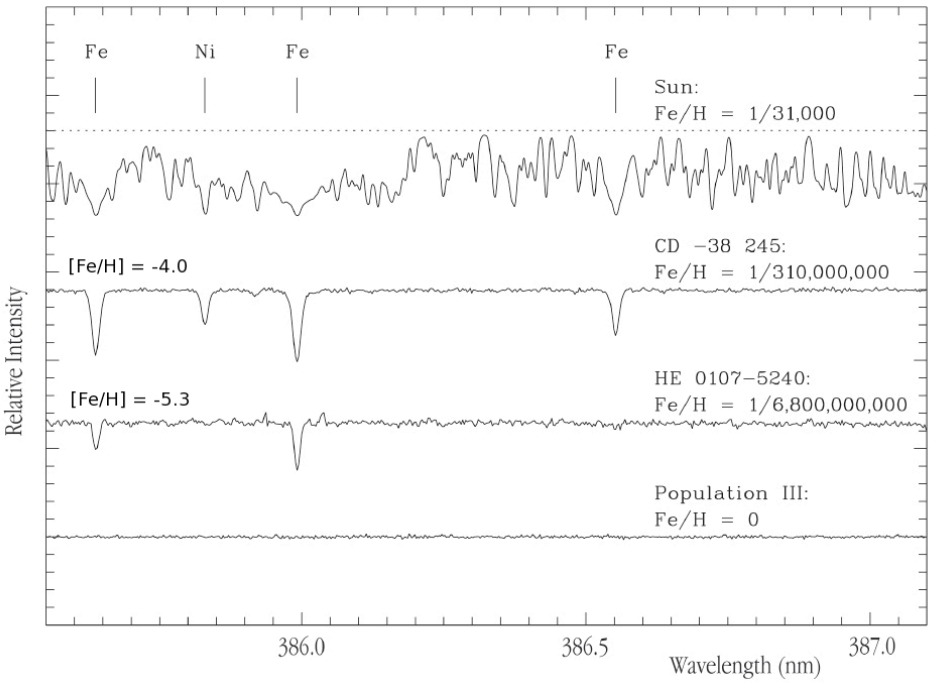


Figure 16.1: *Optical spectra of the sun and the hyper metal poor stars CD -38 245 ([Fe/H] = -4.0) and HE 0107-5240 ([Fe/H] = -5.3). Clearly visible is that the strength of metal lines decreases with decreasing metal abundance. In comparison to CD -38 245 the iron lines in the spectrum of HE 0107-5240 are weaker (or absent) and nitrogen is not visible at all. The bottom spectrum is for a hypothetical Population III star with a chemical abundance pattern as produced in the Big Bang, i.e. H, He and traces of Li.*

accurate, albeit lower abundance $N_{\rm He}/N_{\rm H} = 8.5 \pm 0.07$ %. The latter number (corresponding to $A_{\rm He} = 10.93$) is the present day solar abundance of helium in the outer convection zone. The difference between these two values is explained by slow element diffusion at the base of the convection zone during the solar lifetime.

Finally, the current solar abundance of lithium is about 160 times lower than is derived from meteoritic data. This is a direct consequence of the low nuclear binding energy of Li, causing the destruction of lithium at temperatures of only a few million kelvins by (in a thermonuclear sense: warm) proton collisions. The low $A_{\rm Li}$ implies that gas is transported between the surface and deeper, hotter layers. This probably occurs in convective motions, though possibly mixing through internal (differential) rotation may also play a role. Li- as well as Be-depletion is observed in F- and G-type stars. Also B kan become depleted through thermonuclear burning in layers that are connected with the surface. The behavior of the boron depletion is, however, less well understood as its only potentially useful spectral line in the optical spectral range is in de violet (B II λ3451), where detectors are not very sensitive.

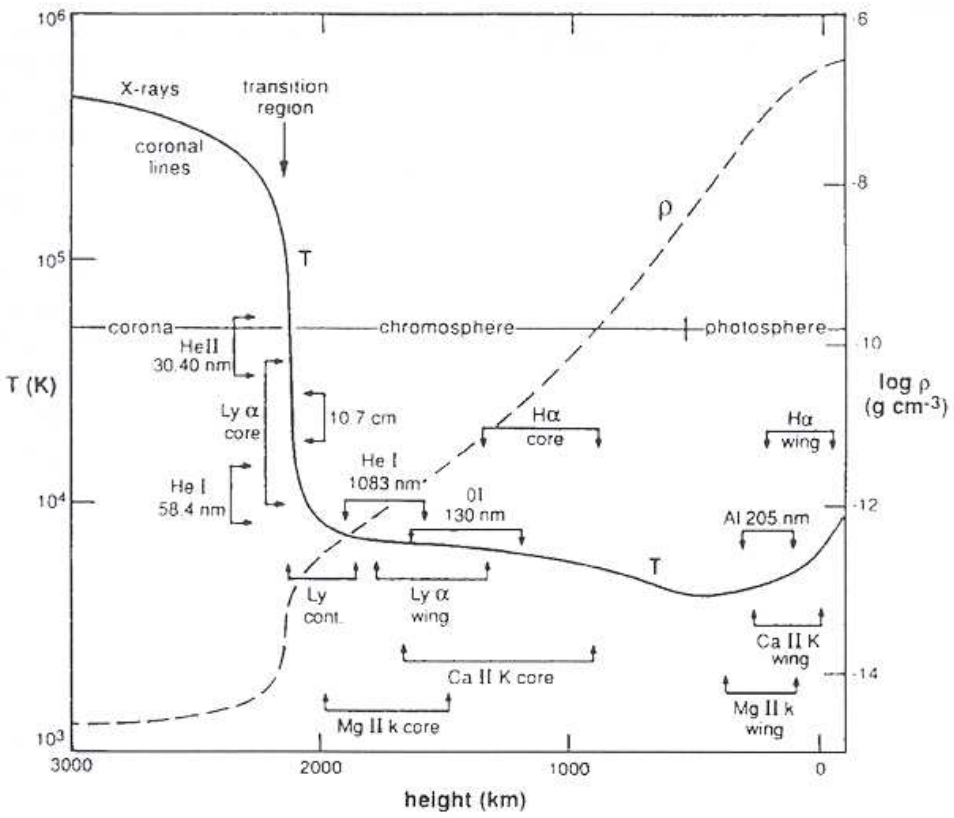


Figure 16.2: *Temperature and density as a function of height for a model of the quiet sun. The zones from which several continua and important spectral lines originate are indicated.*

Structure of the atmosphere

The solar atmosphere can be divided in several zones (see figure 16.2). From inside out these are: (1) the photosphere, from which the visual spectrum originates and where the temperature decreases with height; (2) a zone around the temperature minimum at about 4 400 to 4 500 K; (3) an extended *chromosphere* in which the temperature is increasing up to values sufficiently high to cause (partial) hydrogen ionization; (4) a small *transition region* separating the chromosphere from the corona, with a very steep temperature gradient. This gradient is relatively large just above the chromosphere but flattens in higher regions; (5) a hot *corona* ($T > 10^6$ K) where hydrogen is fully ionized.

Suppose, we secure the spectrum of a quiet spot on the solar surface. As the extinction is wavelength dependent, we can, by observing from X-ray through radio, probe different heights in the solar atmosphere. For a wavelength at which the extinction is small we probe radiation originating from a deep layer. Likewise, for a wavelength at which the extinction is large we see a layer that is higher up. At 1.6 μm, i.e. the wavelength at which the H^- extinction has

a local minimum (see § 8.3 and the upper panel of figure 8.6), we can see deepest into the sun. At the center of the solar disk the brightness temperature of this layer is 6 800 K. The disk averaged T_b at 1.6 μm is about 500 K lower due to limb darkening. At larger (infrared) wavelengths the extinction increases implying that the radiation that we observe originates from cooler layers. The brightness temperature reaches a minimum of about 4 500 K at $\lambda = 150\,\mu\mathrm{m}$. Beyond this wavelength the extinction still increases, but the observed T_b increases due to the increasing chromospheric temperature. At wavelengths shorter than 1.6 μm we see a decrease in the brightness temperature because of an increasing extinction. In the ultraviolet at about 1600 Å T_b reaches a minimum of 4400 K. The extinction continues to increase for even shorter wavelengths, but the brightness temperature runs up due to the chromospheric temperature increase.

Spectral lines also contain information on the temperature structure in the solar atmosphere as the extinction at line center is much larger than it is in the wings or in the continuum next to the line. For instance, the strong Ca II resonance lines at λ3934 (Fraunhofer's K line) and λ3968 Å (H line) have two small emission peaks at both sides of line center because of the temperature increase in the chromosphere (see figure 16.3). These peaks are observed to vary in strength as a function of time and spatial position on the solar disk and are sensitive to local velocity fields.

The temperature rise in the upper layers of the chromosphere and in the transition region between chromosphere and corona is the result of a competition between two effects: a mechanical heating and a strong cooling by resonance lines of hydrogen and helium. This competition results in a flattening of the temperature in the uppermost layers of the chromosphere at about 8 000 K, such that hydrogen is only partly ionized. In even higher, more rarefied layers hydrogen gets fully ionized (see eq. 6.24 and realize that n_e is decreasing further) and the line cooling mechanism becomes inefficient. Consequently, the temperature shows a steep rise. The amount of energy needed to let the temperature increase with distance is modest: only $\sim 10^{-4}$ of the total luminosity of the sun is needed to heat the chromosphere. The energy that is needed to heat the corona – which is even less dense – is extremely modest. It is an order of magnitude less than that needed for the chromosphere, and the largest part of it is transported back to the upper part of the transition region by means of thermal conduction, where it is radiated away by H en He resonance lines.

Heating of chromosphere and corona

What mechanism heats the chromosphere and corona? Though solar physicists have been investigating this seemingly simple question for several decades the answer is still poorly known. Concerning the chromospheric heating some consensus appears to have been reached: acoustic waves originating in turbulent convective layers below the photosphere could at least heat the lower regions of the chromosphere. When these waves propagate outward through the atmosphere, to regions of lower density, the pressure fluctuations grow and shock waves develop. These shocks can heat the gas locally. In the sun, however, the dissipation of these shock waves is so strong that they can not reach the upper regions of the chromosphere. In

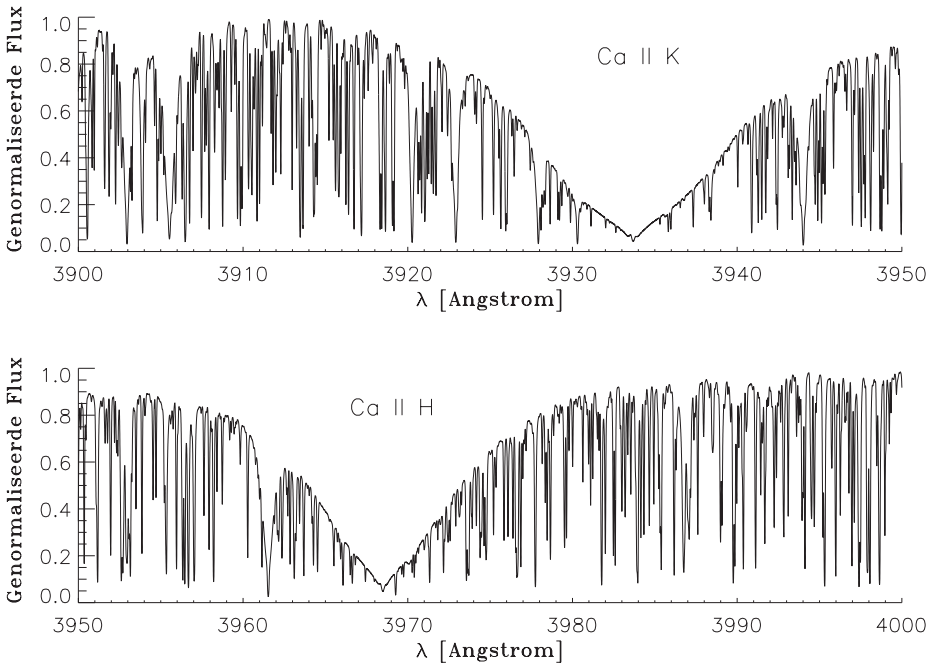


Figure 16.3: *The Ca* II *K in the solar spectrum. This resonance transition, by Fraunhofer refered to as the 'K' line, is the strongest line in the visual part of the solar spectrum. Many weak spectral lines are superposed on the broad wings of the line (so called* blends*); most are from neutral metals such as Fe* I*. In its line core the K line shows two minuscule peaks. The K line is just sufficiently strong for the line source function to feel the onset of the temperature rise towards the chromosphere before other effects start to dominate that cause a decoupling from the local Planck function.*

other stars, notably in M dwarfs, circumstances can be significantly better and acoustic waves may 'survive' into the upper chromosphere. In M dwarfs acoustic waves (of the longest wavelengths) may even heat part of the lower coronal regions. Though our knowledge of coronal heating is still limited, it seems logical to expect that magnetic processes play a dominant role. A possible mechanism may be waves propagating outward along magnetic field lines.

16.3 Solar spectrum

Irradiance at the top of the earth atmosphere

The mean total solar irradiance $\mathcal{R} = (1.365 \pm 0.002)\, 10^6$ erg cm^{-2} sec^{-1} or 1365 ± 2 W m^{-2}. This number used to be called the solar constant, however, as we now know that $\mathcal{R}$ is not constant the term *total irradiance* is prefered. The mean spectral energy distribution of this

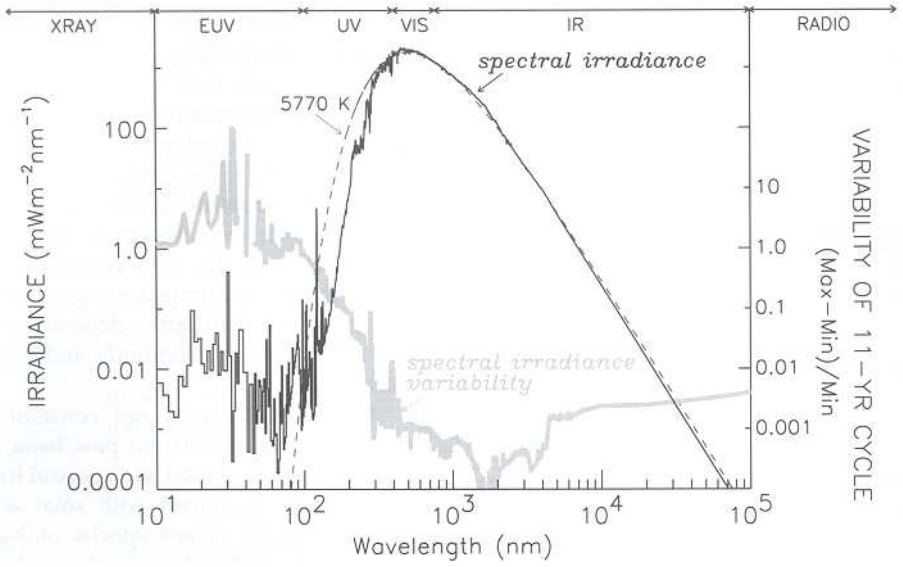


Figure 16.4: *The spectrum of the solar irradiance as observed from above the earth atmosphere. The gray curve provides an estimate of the natural variation of the irradiance during a recent 11 year solar cycle. The variation in the total irradiance is at most 0.2 percent. At 1.6 μm (or 1600 nm) one can see deepest into the solar atmosphere; these layers are the most stable, hence fluctuations in $\mathcal{R}_\nu$ are the smallest. The EUV emission ($\lambda < 912$ Å) is due to non-thermal processes in the chromosphere and corona; at these wavelengths the irradiance shows the largest variations.*

light, i.e. the *irradiance* (see § 3.4), is shown in figure 16.4.

For different reasons the total and monochromatic irradiance fluctuate. Obviously there is a variation as a result of the elliptical path of the earth around the sun causing the distance earth-sun to vary between 1.0167 and 0.9833 AU. This effect, causing a 3% fluctuation on $\mathcal{R}$, is not intrinsic but geometric in nature and officially is not included in the variation of $\mathcal{R}$ or $\mathcal{R}_\nu$. Precise measurements of the (total) irradiance show that fluctuations occur that reach 0.2% on timescales of days to weeks, mainly due to changes in the number of dark spots and bright faculae on the part of the solar surface that we observe. These variations of 0.2% are recorded in visual and near-infrared light, i.e. at the wavelengths where the sun emits most of its energy; for *x-ray* and radio emission such short period variations can lead to fluctuations exceeding a factor of two. During recent epochs of relatively high solar activity, near maxima in the 11 year solar cycle, the mean level of the total irradiance was about 0.1 to 0.15% higher than during cycle minima. A comparison of these relatively recent data to historical observations of the sun (for instance of sunspots), geomagnetic activity and cosmogenic isotopes (e.g. from ^{13}Be in ice-cores and ^{14}C in growth rings of trees) shows that the total irradiance has increased by a few tenth of a percent (between 0.2 and 0.5%, depending on the method used) relative to an epoch of anomalously low solar activity in the seventeenth century, named the *Maunder minimum*. This epoch of exceptional low solar activity is possibly only the most recent one in a series of comparable periods the sun has experienced in previous

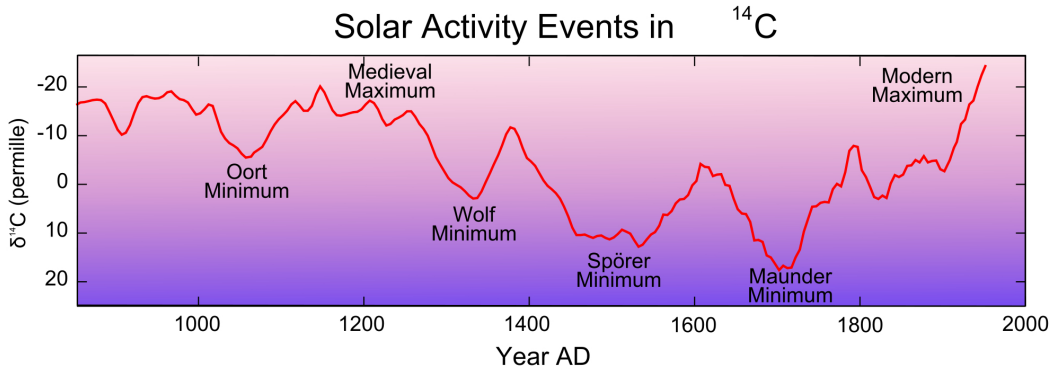


Figure 16.5: *Reconstruction of solar activity events from changes in the ^{14}C isotoop abundance relative to an unspecified calibration. The last datapoint is for 1950. Several periods of low solar activity are labeled. Figure © Leland Mcinnes.*

millennia (see figure 16.5). The Maunder minimum caused a noticeable change in the earth climate. For a thorough understanding of antropogenic (originating in human activity) effects on our climate it is essential to correct for this type of natural variation.

Irradiance at the earth surface

Figure 16.6 compares the irradiance at the top of the atmosphere (light grey / yellow) and that of direct sunlight at sea level (dark grey / red). The spectrum at sea level is significantly attenuated by absorption by the constituents of the earth atmosphere, primarily molecular bands due to water, oxygen and ozone. Water vapor in the atmosphere absorbs much of the radiation past $1\ \mu\mathrm{m}$ via molecular transitions in H_2O. However, water also allows a considerable amount of solar radiation to be transmitted through the atmosphere in certain 'windows' where it has inefficient absorption. Compare these windows to Fig. 2.10, showing photometric transmission curves and identify the location of the J, H and K bands.

Another interesting feature is the difference in radiation above and below the atmosphere on the UV-blue side, where the Wien part of the spectrum starts. This is a combination of the *ozone effect*, as O_3 prevents the life-threatening UV radiation from reaching the surface of our planet, and Rayleigh scattering (see section 8.4).

16.4 Solar activity

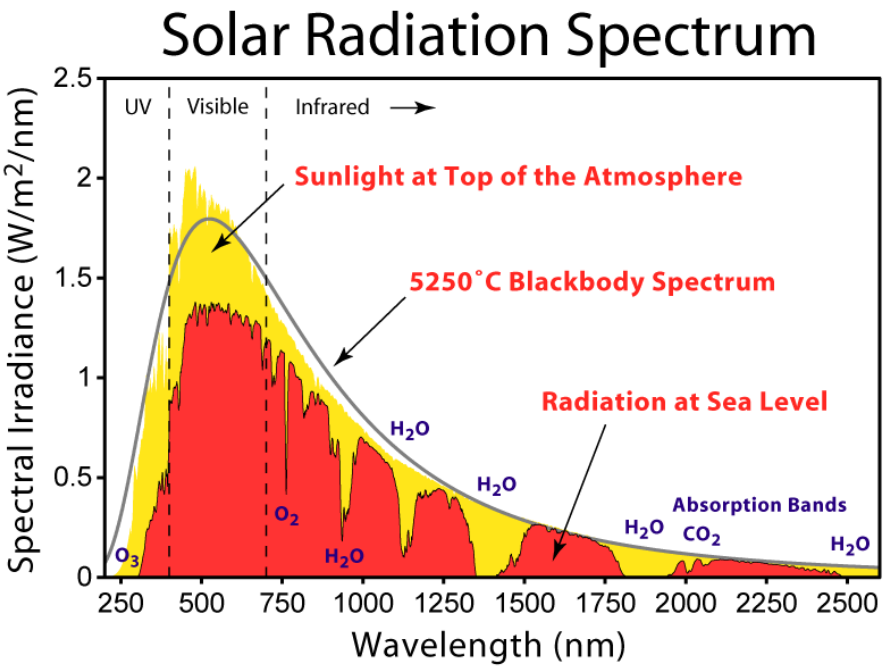


Figure 16.6: *Sunlight as received at the top of the Earth atmosphere (yellow area) and at sea level after attenuation by atmospheric constituents, primarily ozone, oxygen, water and carbon dioxide (red area). The radiation measured at sea level is that of direct sunlight (so not that of diffuse sunlight). The yellow area is the same as that under the black body spectrum, i.e., they represent the same effective temperature. Figure from Robert A. Rhode, http://globalwarmingart.com/wiki/Image:Solar_Spectrum.png.*

Active regions

At almost every moment the part of the solar surface that can be observed from earth shows a number of *active regions*. Near the minimum of the 11 year solar cycle there are only very few, sometimes even none; in the solar maximum there can be as much as five to ten. The structure of these active regions is extremely complex, and we suffice by providing only a general description. Active regions form when the top part of loops of magnetic flux in the solar interior reach into and above the solar atmosphere, where they (thus) can be observed. The shape of such a loop is somewhat similar to the Greek letter Ω. It is thought that the magnetic loops are generated by the solar dynamo, deep in the solar interior.

The shape and appearance of an active region depends sensitively on the observed wavelength and the angle at which it is seen. To give an example: in an active region near the solar rim, seen against the dark sky background, the loop-shaped structure of the magnetic fields can be seen clearly in *x-ray* line emission of highly ionized gas streaming along the field lines. These loops of intensely heated gas are refered to as *coronal arcs* or *coronal loops* (see figure 16.7). High resolution x-ray imaging of active regions near the center of the solar disk show a fine-

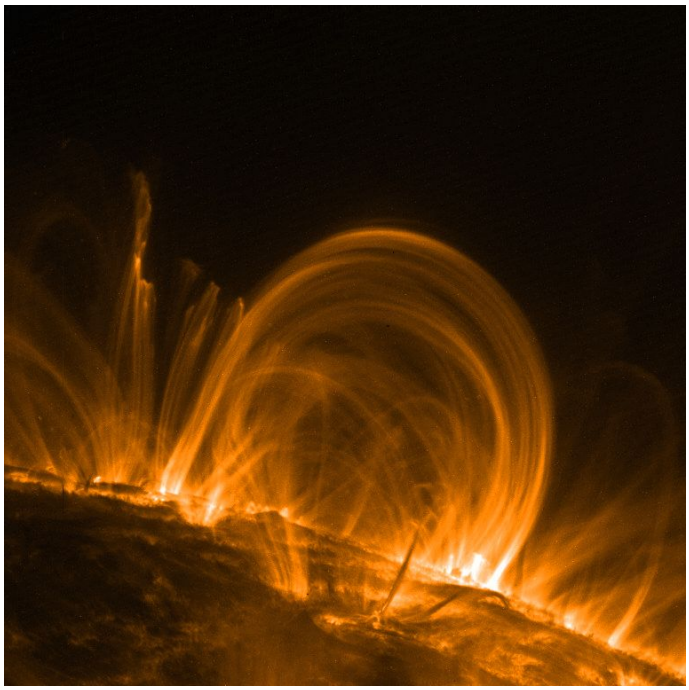


Figure 16.7: *An active region at the edge of the solar disk, observed in a spectral line of highly ionized iron. The temperature of the emitting gas is about 1.5 MK. Photo* © NASA TRACE *satellite.*

structured pattern of magnetic strings connecting the point where the beam of magnetic flux breaks through the photosphere and chromosphere (the north pole) to the point where the field again enters the sun (the south pole). These points are called the *footpoints* and each active region consists of a collection of such bi-polar footpoints, connected by coronal loops. The high density of magnetic field lines in footpoints is what sets active regions apart from quiet regions on the solar surface. The effect of these field lines on the (local) temperature and density structure depends on *a)* the strength of the magnetic field, *b)* the fraction of magnetized relative to unmagnetized gas in the footpoints, and *c)* the topology of the area, i.e. the way in which each point (inside a footpoint) is connected to another point. Regions in which the magnetic field has a 10 to 100 times larger density compared to the surrounding quiescent regions are referred to as *plage* (see figure 16.8). In $\mathrm{H}\alpha$ we see the chromospheric emission from this plage (at temperatures of about 10 000 K) as bright regions, provided that one does not observe it at a too high spatial resolution, as in that case individual chromospheric arcs can be seen. In *sunspots* the field lines are so closely packed that there are (almost) no unmagnetized regions left. This inhibits energy transport through convection, explaining why sunspots appear cooler and darker than their surroundings. Figure 16.8 also shows irregular dark strings or *filaments*. These filaments mark regions of opposite polarity on the solar surface. At a considerable height (in the corona) the field lines will arc over these filaments, creating a sort of tunnel. In this tunnel, along the dividing line of magnetic polarity, referred to as the *filament channel*, magnetic field lines cross the channel allowing clouds of ionized gas to be captured

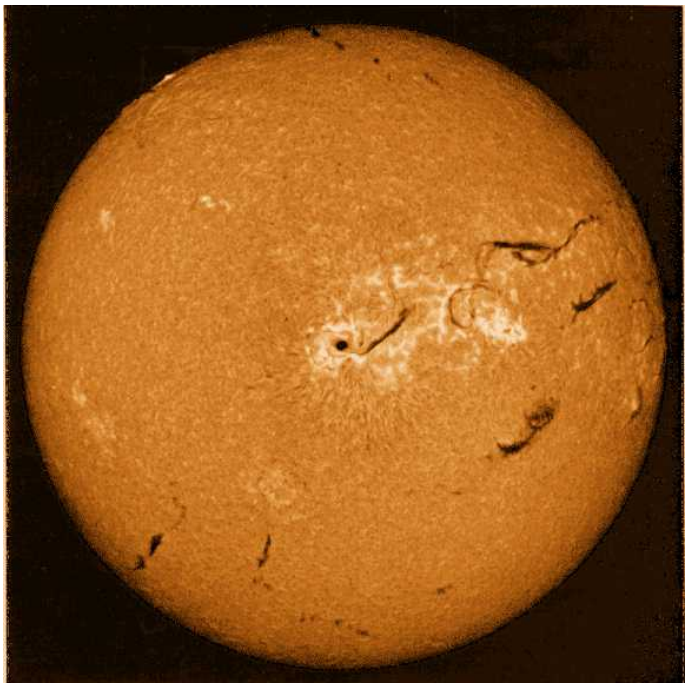


Figure 16.8: *An image of the solar disk in* Hα*, identifying gas at a temperature of about 10 000 K. An extended active region is visible somewhat to the right of the center of the disk, with sunspots, plage, and filaments. Photograph* © the National Solar Observatory.

above the photosphere. These clouds emit in continuum light, but not in Hα or Ca II (hence, appear as dark lines in an Hα image, such as Fig. 16.8). Filaments that are seen in profile at the edge of the solar disk are called *prominences*. Prominences are somewhat similar to coronal loops, but do not show a clear arc structure (they are clouds).

Active regions come and go, with a characteristic lifetime that depends on their size. There is no preferential size for an active region, but they do show a continuous distribution in their sizes. The smallest can barely be resolved with a telescope (and probably this implies that smaller regions exist that we can not detect with the current arsenal of instruments); the largest ones occupy 1–2 % (or 50 000 to 100 000 Mkm^2) of the total solar surface. It is typically these large regions where sunspots and solar flares are seen. The lifetime of the large regions is roughly two months, i.e. one to two rotation periods. The smaller regions have shorter lifetimes.

Solar flares and coronal mass ejections

Like most dynamic phenomena on the sun, the occurrence of solar flares is closely related to the presence and evolution of solar magnetic fields. As a simple introduction (which is all we will do here), we compare flares with earthquakes. The occurrence of earthquakes involves

two steps: energy build-up and energy release. The stress and energy are build up by relative motions of two tectonic plates along fault lines. When the stress reaches a critical point, two plates cannot slide further, the equilibrium brakes down and a section of the fault line is restored to its original position of minimum energy. The excess energy is released in the form of kinetic energy and is propagated away from the epicenter as seismic waves.

A flare can be described by a very similar scenario. We start with two magnetic regions of opposite polarities, such as for instance is marked by filament structure. The line dividing two polarities is called the *magnetic neutral line*. In the starting- or *potential configuration* the field lines run perpendicular to the magnetic neutral line. If, for some reason, magnetic regions slide, the field lines will start to move away from perpendicular to the neutral line. This is called the sheared configuration. In an extremely sheared configuration, the magnetic field lines are nearly parallel to the neutral line. A sheared magnetic field has more magnetic energy than the corresponding potential field configuration and the extra energy is called the *free magnetic energy*. When the fields are extremely sheared, an instability may occur. Fields tend to be restored to the potential configuration and the extra energy is released by the process of *magnetic reconnection*. The free magnetic energy is converted to thermal and non-thermal energy. This causes a solar flare, in which thermal radiation is emitted and non-thermal particles are accelerated.

Solar flares are wonderfully complex phenomena and are observed across the electromagnetic spectrum as well as via energetic particles in space. They are classified according to their size, duration, morphology or magnetic topology and the composition of their associated energetic particles. Despite the seemingly infinite variety in solar flare characteristics, one may identify two basic types – impulsive and gradual. *Impulsive flares* have time scales of the order of a few minutes or less while gradual flare durations range from tens of minutes to several hours. These two basic types are combined in fully developed flares in which an impulsive phase is followed by a gradual main phase. Long-duration gradual flares are also characterized by *coronal mass ejections*.

A coronal mass ejection can carry 10^{15} to 10^{16} gram of plasma into space. The source region within the sun's atmosphere can be hundreds of thousands of km across and the resulting ejection can expand into space at many hundreds of $\mathrm{km\,sec^{-1}}$. The rate of coronal mass ejections varies with the solar cycle; one may expect of the order of 0.1 event per day during solar minimum, upto more than one a day during maximum. They represent a very significant disturbance of the solar wind. Given their size and mass, combined with the fact that the expanding clouds carry a frozen-in magnetic field, such events can engulf the Earth system and their arrival at the Earth can generate significant geomagnetic storms.

Solar cycle and sunspot cycle

Important properties of active regions vary with the 22 year *solar cycle* or the 11-year *sunspot cycle*. Only few active regions, together with the sunspots they contain, are visible during the solar minimum – sometimes (for a few weeks) none are seen at all. During a sunspot

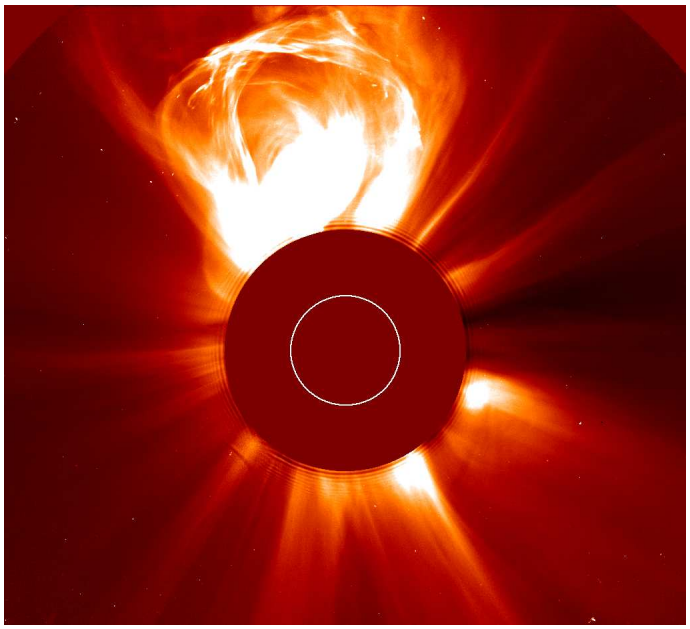


Figure 16.9: *Coronal mass ejection observed by the SOlar and Heliosphere Observatory (SOHO) on March 9 2004. Direct sun light is blocked; the position of the solar disk is given by the half circle. Foto* © SOHO *satellite.*

maximum typically five and sometimes $\sim$ ten active regions are seen from earth and the mean level of the irradiance is found to be about 0.1 to 0.15% higher than during cycle minima.

At the start of a sunspot cycle spots appear at about 30° (sometimes even 40°) north and south of the equator. The sunspot occurrence locations slowly migrate towards the equator. During maxima, about 4 to 5 years later, spots appear in a strip with an average latitude of about 15°. Near the end of the cycle, close to minimum, the strip has narrowed and is centered around 8° and new sunspots start to appear at $\sim$ 30°. In every bi-polar pair of sunspots the one that is in front – in the sense of the rotation direction of the sun – is called the *leader*, and its trailing companion the *follower*. During an entire cycle, from minimum to minimum, all leaders on the northern hemisphere have equal polarity (say $+$) as do all leaders on the southern hemisphere, though the polarity of these southern leaders will be opposite to that of the leaders on the northern hemisphere (in this case $-$). At the end of each cycle the pattern of polarity is reversed (so, in our example, $+$ to $-$ on the northern hemisphere and $-$ to $+$ on the southern hemisphere). A complete magnetic cycle thus consists of two solar cycles of 11 year, and lasts 22 years.

The sunspot cycle is part of a larger wider-ranging cycle in which all aspects of the solar activity, including sunspots, plages, prominences, solar flares and coronal mass ejections, as well as the shape, size and structure of the chromosphere and corona, show a cyclic variability with a period of 11 years. The number and the dimensions and energies of prominences, flares

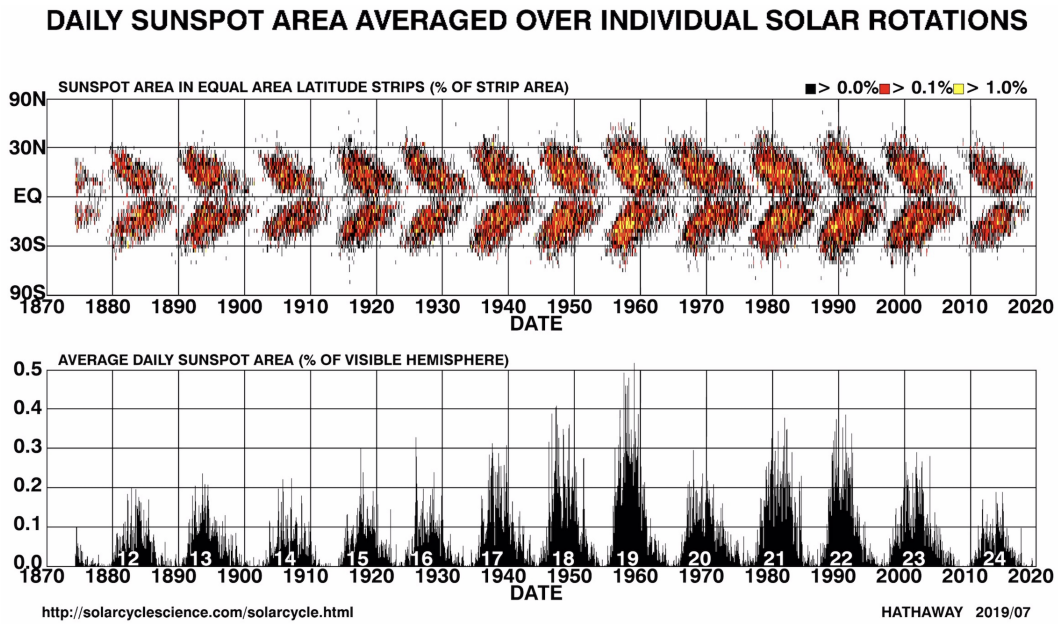


Figure 16.10: *Top: the latitude of sunspots as a function of time. At the start of the 11 year cycle sunspot appear at the circles of latitude at* $+30^\circ$ *and* -30° *distance from the equator. At maximum solar activity, about 4 to 5 year later, sunspots also emerge at the equator. After 11 years the last sunspots of the cycle form near the equator while new ones again appear at 30°, starting a new cycle. The distribution of spots appears symmetrical around the equator and is known as the* butterfly diagram. *Bottom: the average percentage of the visible solar surface that is covered by sunspots as a function of time. Notice that the current cycle (cycle 24) appears not to develop itself as prominently as previous cycles. Source: http://solarcyclescience.com.*

and coronal mass ejections reflect the increase and decrease of the number of sunspots, the corona being brighter, more extended and more symmetric during minimum activity.

Solar wind

The solar wind is a continuous, more or less spherical outward flow of charged particles – mostly protons and electrons – from the sun into interplanetary space. The outflow "blows" past the planets at speeds that fluctuate between 200 and 900 $\mathrm{km\,sec^{-1}}$, and cause the sun to loose about $2 \times 10^{-14} M_\odot$ of mass per year. At at distance of 1 AU from the sun, the mean density of the solar wind is about 5 particles per cm^3 and its temperature is about 10^5 K. Close to the plane of the ecliptic, the average wind speed is about 400 $\mathrm{km\,sec^{-1}}$. The particles that comprise this relatively slow wind emerge predominantly from streamer-like structures in the solar corona that tend to be located relatively close to the solar equator. A relatively fast wind, with a mean speed of about 750 $\mathrm{km\,sec^{-1}}$, flows out of coronal holes that are centered permanently on the solar poles. When coronal holes (temporarily) extend down to, or across, the solar equator, high-speed streams flow out past the Earth and planets (such that the wind

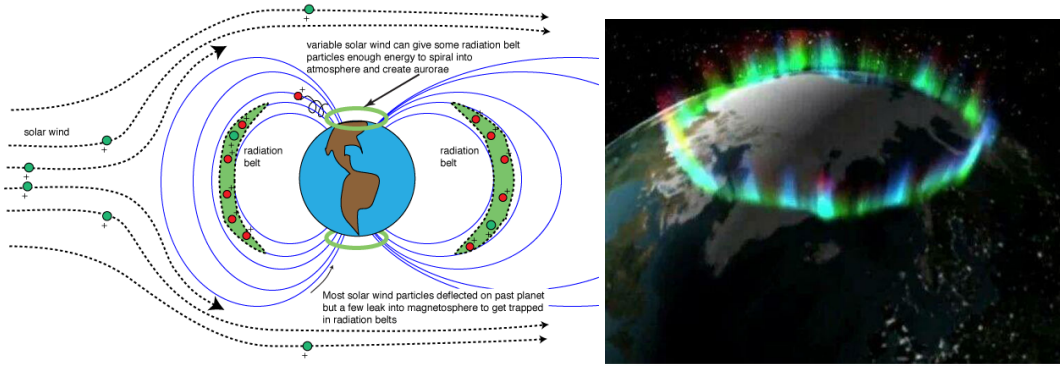


Figure 16.11: *Left: Schematic picture of ionized solar wind particles streaming past Earth. Some are trapped in the Earth magnetosphere in so called* radiation belts*. Variations in the strength of the solar wind can give some radiation belt particles enough energy to spiral into the atmosphere along the Earth magnetic field and create aurorae in the northern and southern polar regions. Right: image from space showing the aurorae borealis. Source: NASA.*

speed shows a period modulation with the solar rotation period). This explains the fluctuation in the solar wind speed.

The solar wind carries with it lines of magnetic force, which spread out to form the weak interplanetary field. The sun's rotation, combined with the radial outward flow of the solar wind, cause the interplanetary field lines to take up a spiral form. The solar wind outflow is eventually stopped when the pressure of the solar wind equals that of the interplanetary medium. This boundary is at about 120 AU and is referred to as the *heliopause*; the enclosed region is the *heliosphere*. The radius of the heliosphere is expected to vary with the solar cycle. The Voyager 1 spacecraft passed the heliopause on August 25, 2012, and became the first manmade object to enter interstellar space.

16.5 Aurora Borealis

The bulk of the solar wind particles impinging on the Earth magnetic field are deflected and stream past our planet. Some, however, leak into the magnetosphere to get trapped in *radiation belts*. Radiation belt particles may spiral into the atmosphere along the Earth magnetic field lines and create aurorae in the northern and southern polar regions (see Fig. 16.11).

When these protons and electrons reach the upper atmosphere they excite (through a collision shower) atoms and molecules. Almost all auroral light consists of emission lines and bands of neutral or ionized N_2, O, O_2, and N. The origin of the most copious green emission was established by McLennan and Schrum in 1925 (see Ireton 1925) as due to the metastable (forbidden) transition of atomic oxygen from the ^{1}S state to ^{1}D state (see Fig. 16.12). The

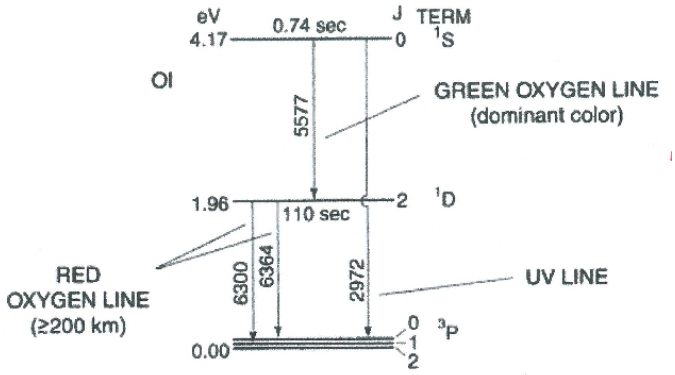


Figure 16.12: *Excited states of the oxygen atom that give rise to forbidden transitions prominent in auroral emissions, and their lifetimes (after Roach and Smith, 1967).*

lifetime of this state is 0.74 s. An even longer-lived transition is the red-line triplet $^1\mathrm{D} \rightarrow {}^3\mathrm{P}$ at 6300−6364 Å with a lifetime of 110 s. This means that the ^{1}D excited atoms can move a significant distance after their excitation and, consequently, the red auroral light is more diffuse than the green light at 5577 Å.

Auroras typically occur in the thermosphere at altitudes between 100 km and 1000 km. The visual aurorae appear in a great variety of forms, including arcs, spirals and curls, and on spatial scales from kilometers to the size of the entire auroral oval with a circumference of the order of 10 000 km. Sometimes the structures remain static for tens of minutes, but often there are rapid variations and motions. These variations are linked to variability of the solar wind – a phenomenon referred to as *space weather*.

The critical density (see eq. 15.14) of the two ^{1}D lines are $n_\mathrm{e}^\mathrm{crit} = 1.4 \times 10^6\,\mathrm{cm}^{-3}$ (for the red λ6300 line) and $4.6 \times 10^5\,\mathrm{cm}^{-3}$ (for the red λ6364 line). These lines are effectively quenched below 200 km where $n_\mathrm{e} > n^\mathrm{crit}$. The critical density of the green line is about two orders of magnitude larger ($7.4 \times 10^7\,\mathrm{cm}^{-3}$), and it can be seen at altitudes as low as about 100 km. The typical energies of the primary auroral electrons range between 1 and 10 keV, which allows them to penetrate to a height range between 150 and 100 km. It is therefore the green light at 5577 Å that dominates the visual appearance of auroral arcs. For very soft electron spectra (for energies $<$ 1 keV, say) the red emissions above 200 km become prominent.

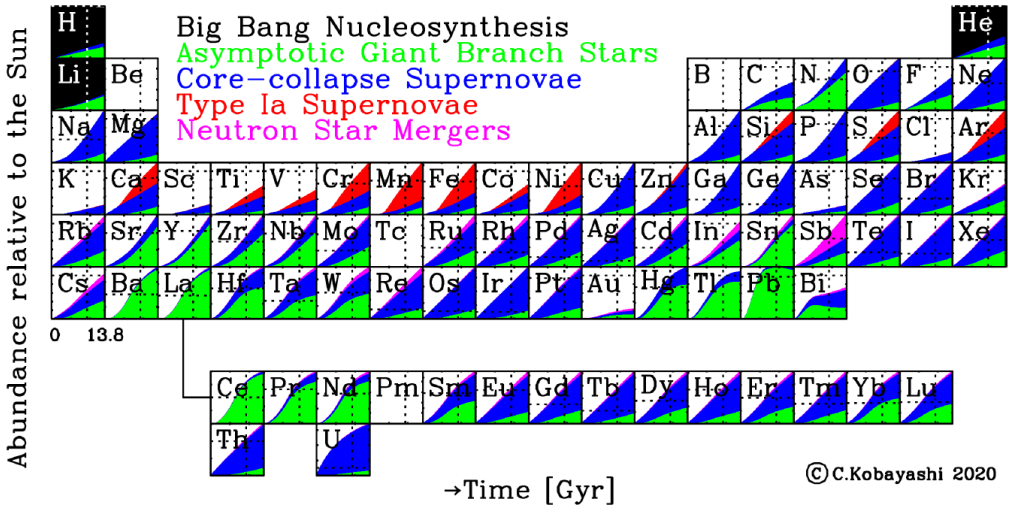


Figure 16.13: *The cosmic production of the elements from Galactic chemical evolution models by Kobayashi et al. (2020). In the background of each element box in the periode table the main producers of the element are shown as a function of cosmic time (see legend). The amounts returned via mass loss are also included for Asymptotic Giant Branch (AGB) stars and stars ending as core-collapse supernovae. The x-axis of each box shows time from $t = 0$ (Big Bang) to 13.9 Gyrs, while the y-axis shows the linear abundance relative to the sun, $X/X_\odot$. The dotted lines indicate the observed solar value, i.e., $X/X_\odot = 1$ and 4.6 Gyrs for the age of the sun (measured back from the present time).*

Exercise 16.1

Determine the mass of the sun from Newton's gravitation law and the centrifugal force. Use the Earth as a test particle and assume that the Earth orbit is circular and the Earth mass to be negligible compared to the solar mass. The gravitational constant is $G = 6.67 \times 10^{-8}$ dyn cm^2 gr^{-2}.

Exercise 16.2

The ten most important elements in life on Earth are H, C, N, O, Na, P, S, Cl, K, and Cl (see *An introduction to Astrobiology*, Revised edition 2011, Cambridge University Press). By percent mass they contribute about 9.5, 18.5, 3.2, 65, 0.2, 1.0, 0.3, 0.2, 0.4, and 1.5 %, respectively, to the human body.

a) Which percentage do these building blocks of life, save H, represent of the total of elements heavier than helium?

b) Are stars 'geared' towards producing the elements necessary for life on Earth or are these elements only a byproduct of stellar nucleosynthesis?

Table 16.3: *Atomic data for transitions in the $2p^4$ configuration of* O I*. The data is compiled by Pradhan & Nahar (2011). The average collisional strength Υ (see eq. 7.12) is given for several temperatures in units of 10^4 K. In the lower parts of the thermosphere, where most of the auroral light originates, $T \sim 200$ K. We estimate that $\Upsilon(0.02) \sim \Upsilon(0.5)/3.4$. For $\Downarrow$, use the number give above $\Downarrow$.*

Ion	Transition	λ (Å)	A_{ul} [s^{-1}]	$\Upsilon_{lu}(0.5)$	$\Upsilon_{lu}(1.0)$	$\Upsilon_{lu}(1.5)$	$\Upsilon_{lu}(2.0)$
O I	$^1D_2 - {}^3P_0$	6393.5	7.23×10^{-7}	1.24×10^{-1}	2.66×10^{-1}	–	5.01×10^{-1}
	$^1D_2 - {}^3P_1$	6363.8	2.11×10^{-3}	$\Downarrow$	$\Downarrow$		$\Downarrow$
	$^1D_2 - {}^3P_2$	6300.3	6.34×10^{-3}	$\Downarrow$	$\Downarrow$		$\Downarrow$
	$^1S_0 - {}^3P_1$	2972.3	7.32×10^{-2}	1.53×10^{-2}	3.24×10^{-2}	–	6.07×10^{-2}
	$^1S_0 - {}^3P_2$	2959.2	2.88×10^{-4}	$\Downarrow$	$\Downarrow$		$\Downarrow$
	$^1S_0 - {}^1D_2$	5577.3	1.22	7.32×10^{-2}	1.05×10^{-1}	–	1.48×10^{-1}
	$^3P_0 - {}^3P_1$	1.46×10^6	1.74×10^{-5}	1.12×10^{-2}	2.65×10^{-2}	–	6.93×10^{-2}
	$^3P_0 - {}^3P_2$	4.41×10^5	1.00×10^{-10}	1.48×10^{-2}	2.92×10^{-2}	–	5.36×10^{-2}
	$^3P_1 - {}^3P_2$	6.32×10^5	8.92×10^{-5}	4.74×10^{-2}	9.87×10^{-2}	–	2.07×10^{-1}

c) Figure 16.13 shows the time evolution (in Gyr) of the origin of the elements in the periodic table. Estimate (roughly) which percentage of the human body (by mass) originates from Big Bang nucleosynthesis, AGB stars, core-collapse Supernovae (so, from massive stars) and Type Ia supernovae (so, from low-mass stars).

Exercise 16.3

The Ca II K line, at λ3933 Å, is the strongest line in the visual spectrum of the sun. At both sides of the line core, within one Angström from line centre, the K line shows two minuscule peaks (see figure 16.3) that have been discussed extensively in the literature. Can you draw four panels as in figure 4.5 that explain these minuscule peaks? Assume that the profile shape of the K line has the standard "clock" shape.

Exercise 16.4

a) Use the data in table 16.3 to compute the critical densities (see eq. 15.14) for the auroral lines of O I at λ5577 Å (green), λ6300 Å and λ6364 Å (both red). Recall that the statistical weight $g_J = 2J + 1$.

17

Stellar winds

A stellar wind is a continuous outflow of matter from a star. This outflow plays an important role in the life cycle of gas and dust in the interstellar medium and in the evolution of stars.

After a short introduction we discuss the most characteristic signature of mass loss in a stellar wind, i.e. the P Cygni profile. From these profiles one can easily determine the terminal velocity of the stellar wind. Next, we focus on several methods to determine the rate of mass loss from stellar winds from spectroscopic and/or photometric observations. For hot stars there are three such approaches that use, respectively: *i)* non saturated P Cygni profiles, *ii)* emission lines, and *iii)* infrared and radio excess.

The most sensitive diagnostics of stellar winds are the P Cygni profiles of resonance lines in the ultraviolet part of the spectrum. Saturated profiles of these lines only yield lower limits to $\dot{M}$. Intrinsically strong resonance lines of stars with mass loss can already display a P Cygni shape if $\dot{M} \sim 10^{-9}\ M_{\odot}\mathrm{yr}^{-1}$. If the line is not saturated one can derive the number density of absorbing particles n_{ijk} as a function of distance from a comparison of the observed profile with model predictions, and by that the mass-loss rate. For resonance lines it holds that $n_{0jk} \simeq N_{jk}$. To accurately derive the mass loss it is therefore necessary that the ionisation properties $q_{jk}(r)$ are known. Sometimes, if it is safe to assume that the relevant ion is dominant throughout the entire wind, i.e. $q_{jk}(r) \simeq 1$, the required computations are relatively simple and a reliable mass loss is found. In most cases the ion responsible for the P Cygni profile being studied is not the dominant ion. Sometimes it is even far from it. In that case there is no other option than to do a proper computation of the state of the gas in the stellar wind. These computations are extremely complex as both NLTE effects, line-blanketing effects, the stellar wind, non-thermal radiation as a result of instabilities in the stellar wind (read: shocks), and structure in the outflow (read: clumps) need to be included in a self-consistent way. This is such a complicated and extensive problem that it is at (or even beyond) what we are currently capable of.

The emission line method for mass loss determinations usually employs the wind emission in the Hα profile (though for very hot stars also He II λ4686 may be a useful diagnostic) and yields reasonably reliable results. The most reliable way to derive $\dot{M}$ is that based on radio flux measurements. In §§ 17.4 and 17.5 we will discuss the principles of these latter two

methods.

For completeness, in the case of cool stars the diagnostics of mass loss and terminal velocity are: *iv)* molecular emission lines, and *v)* infrared and (sub-)millimeter continuum radiation by dust that condenses in the stellar wind outflow. These methods will not be discussed in these lecture notes.

17.1 Historical introduction

The P Cygni profiles that are so characteristic for stars with a stellar wind were, you never guess, first seen in the star P Cygni. This star – to provide a nice trivia – was discovered as a "new star" in the summer of 1600 by the Dutch chartographer Willem Janszoon Blaeu (1571 – 1638). The discovery of the characteristic profiles and insight in their meaning followed only much later.

Beals (1929) noticed that both the spectra of Wolf-Rayet stars (see § 2.2) and of novae showed P Cygni profiles. Photographic plates of novae, first that of Nova Aurigae in 1891, showed that these events were associated with the formation of shells of gas centered around the star that experienced the outburst. Continued observations of these shells, over periods of years, showed that the diameters of the shells increased, associating P Cygni profiles with the outward expansion of gaseous envelopes. An important difference between Wolf-Rayet stars and novae was that the P Cygni profiles in the WR-stars do not change over time. This led Beals to propose that a continuous outflow occurs from Wolf-Rayet stars. This was confirmed by Chandrasekhar (1934), who developed a solid footing for interpreting P Cygni profiles as arising in expanding atmospheres (see § 17.2). Korisev (1934) used the diagnostics developed by Chandrasekhar to estimate the mass loss and maximum outflow velocity of a Wolf-Rayet star and found, respectively, $\sim 10^{-5}\ M_{\odot}\mathrm{yr}^{-1}$ and $\sim 1000\ \mathrm{km\,sec}^{-1}$.

In the case of cool stars, it was more difficult to find evidence for mass loss. The crucial evidence was presented by Deutsch (1956), who showed that the M5 giant in the binary system α Hercules looses mass at a rate of $\sim\ 10^{-7}\ M_{\odot}\mathrm{yr}^{-1}$ and at a maximum flow speed that reached $\sim\ 10\ \mathrm{km\,sec}^{-1}$. Deutsch noticed that the spectrum of the accompanying G5 star contained a set of shifted lines that are also present in that of the M star. He realized that the orbit of the G-dwarf is embedded in the expanding envelope of the M giant, making the G-dwarf serve as a "probe" of the velocity and density in the remoter parts the M giant outflow. As the measured velocity in these remote parts of the M giant atmosphere was in excess of the local escape speed he concluded that the gas would escape from the M giant.

In addition to spectroscopic evidence, Wilson (1960) realized from stellar structure considerations that stellar mass loss must occur. For stars with initial masses in excess of the Chandrasekhar limit of 1.4 $M_{\odot}$, it was necessary that mass was lost in order for these objects to evolve to the white dwarf phase.

Research into the properties and nature of stellar winds expanded dramatically in the 1960s. In 1962 the existence of a solar wind was observationally confirmed by the *Mariner 2* interplanetary probe. In 1967, Morton, using a balloon experiment, secured UV spectra of several O and B supergiants. These showed P Cygni profiles, a telltale signature of stellar winds. Since that time many observations, ground based as well as from rockets, balloons and satellites, have conclusively shown that stellar winds occur in many different types of stars. The evidence is found in almost all wavelength domains, from X-ray to extreme-UV, UV, optical, infrared, sub-millimeter to the radio regime. Space observatories that have contributed very significantly to these studies include the *Copernicus* satellite, the *International Ultraviolet Explorer* (IUE), the *Einstein* röntgen observatory, the Dutch-American *Infrared Astronomy Satellite* (IRAS), *Extreme Ultraviolet Explorer* (EUVE), *Hubble Space Telescope* (HST), the partly Dutch *Infrared Space Observatory* (ISO), and the *Herschel Space Observatory*.

17.2 P Cygni profiles

In general one can easily distinguish spectral lines that are formed in a stellar wind from those that orginate in the photosphere. The wind lines have a much larger width, a result of the acceleration of the outflowing gas, compared to the photospheric lines that are formed in the (more or less) hydrostatic parts of the atmosphere and which are only affected by thermal and turbulent broadening. The wind lines have either absorption or emission profiles, or a combination of these two: the so called P Cygni profile.

P Cygni profiles usually occur if the line forms through the process of *resonance scattering*. This occurs if the transition is between the groundlevel (or a meta-stable level) and an excited level. Often, the upper level is the first excited level, however this is not strictly required. In a scattering the extinction of a photon in a photo excitation is followed by the spontaneous emission of a photon of (virtually) equal frequency. The photon is therefore not lost, but has only changed direction. Examples of resonance lines are N V $\lambda\lambda$1239,1243, C IV $\lambda\lambda$1548,1551, and Si IV $\lambda\lambda$1394,1403 in the spectra of O and early B-type stars; C II $\lambda\lambda$1335,1336 in late B- and A-type stars, and Mg II $\lambda\lambda$2796,2803 in late B- to M-type stars. Notice that these are all lines of relatively abundant elements. Of course, Lyα λ1216 will be strong in all these types of stars.

Schematic explanation of the formation of a P Cygni profile

Figure 17.1 provides a schematic representation of the formation of a P Cygni profile. We assume a spherically symmetric outflow in which the velocity of the gas increases monotonically, reaching a *maximum outflow velocity* or *terminal velocity* v_∞. A distant observer can identify four regions that each contribute in a specific way to the line profile.

a) The photosphere of the star, emitting the continuum "on top of which" the line is

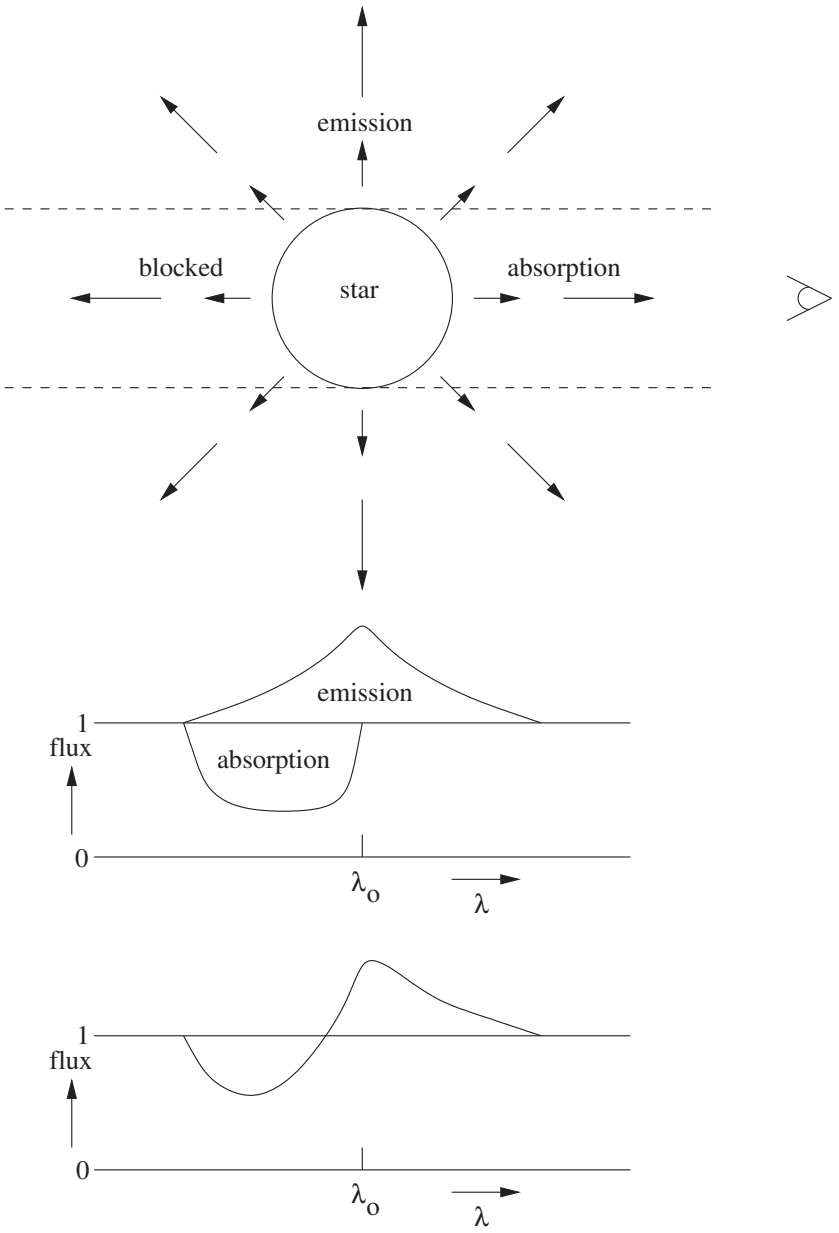


Figure 17.1: *Schematic representation of the formation of a P Cygni profile. The observer can identify four regions: a) the photosphere of the star, from where the continuum originates; b) the 'tube' in front of the stellar disk, where blue shifted absorption occurs; c) the 'tube' behind the stellar disk, which can not be observed, and d) the halo surrounding the stellar disk from which the symmetrical emission profile originates. In a fully developed P Cygni line the contributions from the regions b and d dominate the profile. In first order, the P Cygni profile is the sum of a, b, and d.*

formed. In almost all cases the photosphere is formed in a layer that has an outflow velocity that is very low compared to the local thermal Doppler velocity, i.e. $v \ll \xi_{\rm D}$. Therefore, a "normal" absorption profile will form at the rest wavelength λ_{lu} of the line. This absorption profile may not be visible as such because the more outward stellar wind may "deform" the profile.

b) A tube like or cylindrical region in front of the stellar disk, where gas is moving in the direction of the observer with velocities that range between $v \simeq 0$ and the maximum outflow velocity $+v_\infty$. In this tube, continuum photons at the blue side of λ_{lu}, i.e. photons with shorter wavelength, will be scattered out of the line of sight towards the observer. If $v_\infty \gg \xi_{\rm D}$ a broad blueshifted absorption profile will form.

c) A tube like or cylindrical region behind the stellar disk, where gas is moving away from the observer. As the stellar disk is occulting this tube, radiation emitted in this region can not be observed. It implies that, at least in principle, the maximum redshift of line photons, i.e. those corresponding to the velocity $v = -v_\infty$, can never be seen. In practice, however, this effect is so small that it can not be identified.

d) The region next to the stellar disk that the observer would see as a halo surrounding the stellar disk if one would be able to spatially resolve the wind outflow. The gas in this halo can have both positive (i.e. blue shifted) as well as negative (i.e. red shifted) velocities relative to the observer and causes a symmetric emission profile with Doppler-shifted velocities between $-v_\infty$ and $+v_\infty$.

In shaping the observed profile the contributions of regions b and d are dominant, *proviso* the stellar wind is not very weak (in which case a will dominate). The net effect of all four contributions – which in a simplistic way of thinking one might view as the sum of four independent profiles – yields the P Cygni profile (see figure 17.1).

The determination of the terminal flow velocity

The maximum outflow velocity can easily be determined from P Cygni profiles that have a strongly saturated blue shifted absorption. To give an example of such saturated P Cygni profiles we show in figure 17.2 a number of UV resonance lines of the O4 If star ζ Puppis. This star at a distance of 450 parsec is the closest early-O supergiant.

If we neglect the thermal velocity $\xi_{\rm D}$ (see eq. 13.16), gas that is approaching us with the terminal velocity will cause the most blue shifted absorption in the line profile. The wavelength of this maximum Doppler shift is, using eq. (13.13)

$$\lambda_{\rm edge} = \lambda_{lu}\,(1 - v_\infty/c) \tag{17.1}$$

If the line is so strong that even at (read: very close to) the terminal velocity all stellar light is extincted there will be no flux at the long wavelength side of $\lambda_{\rm edge}$, i.e. the line is saturated,

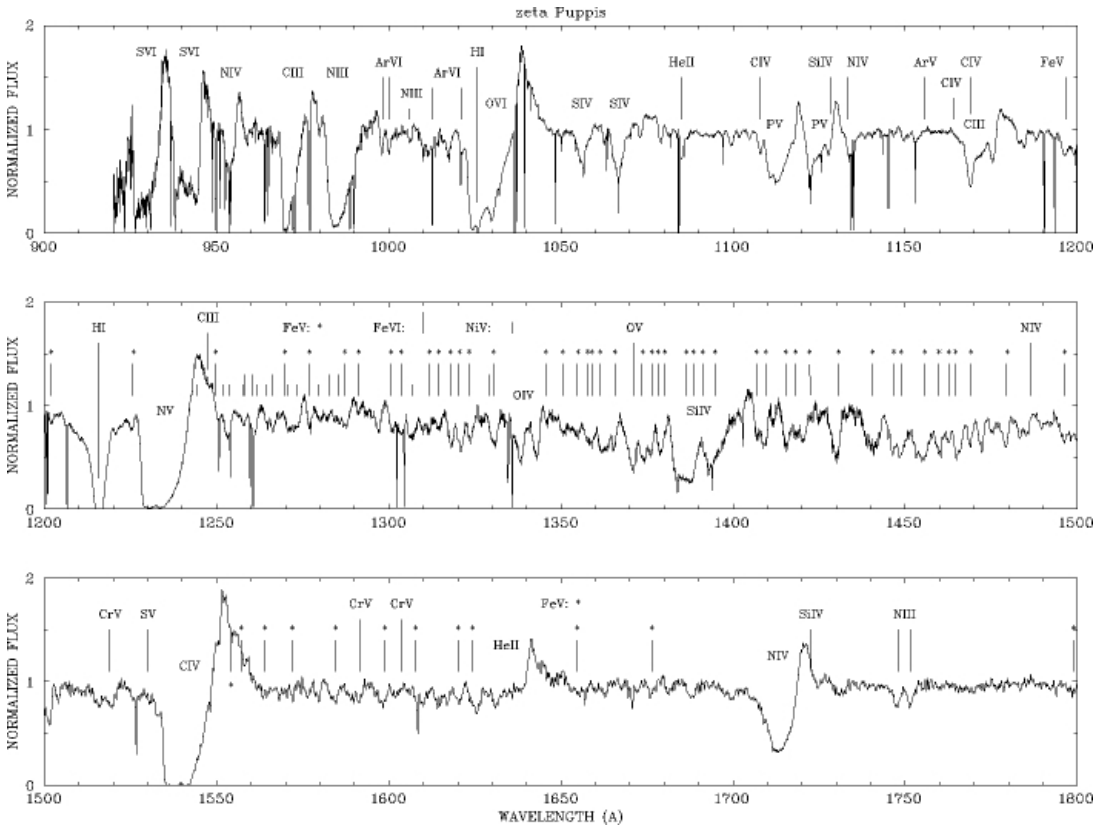


Figure 17.2: *Merged spectrum of Copernicus and IUE ultraviolet high-resolution observations of the O4(f) supergiant ζ-Puppis. The most important wind lines of the light elements are identified and marked. Also marked are the large number of lines from iron group elements (e.g. Fe* V*) which are especially present between 1250 and 1500 Å. Most of these iron-group lines are formed in the photosphere* and *in the lower part of the outflow. The resonance P Cygni profiles of O* VI*, N* V *and C* IV *are saturated and can be used to measure the terminal velocity. Figure from Pauldrach et al. (1994).*

while suddenly at the short wavelength side the stellar continuum will be visible again. By simply measuring the wavelength at this abrupt transfer from no flux to the full continuum flux, the terminal velocity is found using eq. (17.1).

Though the above is a trivial procedure, there are a number of effects that can complicate the determination of v_∞. These are: *a)* The P-Cygni profile is not saturated. In this case the bluest wavelength showing line extiction no longer corresponds to that of the terminal velocity, and only a lower limit for v_∞ can be measured; *b)* The ionisation of the ion responsible for the P Cygni profile may at some point in the outflow suddenly decrease. If this happens, the same scenario as in case *a* applies, i.e. one measures a lower limit for the terminal velocity; *c)* The stellar wind is turbulent. Both observations and theory point to the presence of strong turbulent motions in stellar winds. At the radial distance where $v \sim v_\infty$ these motions can reach an amplitude $\xi_{\rm micro} \sim 0.05 - 0.15\, v_\infty$. If this turbulence is not taken into account the

terminal velocity will be overestimated by roughly 10 to 30 percent; *d)* At the blue side of the P-Cygni profile, just past the wavelength corresponding to v_∞, some absorption lines of other elements may be present. If one would not be aware of the presence of such lines, one would overestimate the terminal velocity.

Velocity law

Observations and theory of stellar winds both show that the behaviour of the outflow speed can be approximated by a *β-type velocity law*, first introduced by Castor and Lamers

$$v(r) = v_\circ + (v_\infty - v_\circ)\left(1 - \frac{R_\star}{r}\right)^\beta \tag{17.2}$$

This law describes a monotonic increase of the outflow velocity with distance from $v_\circ$ at the photosphere ($r = R_\star$) to a maximum velocity v_∞ at large distance. For O- and B-type stars the starting velocity is subsonic, whereas the terminal velocity v_∞ is highly supersonic. It actually holds that $v_\circ \ll v_\infty$. The parameter β is a measure of the velocity gradient or "steepness" of the velocity law. For early-type stars $\beta \simeq 1$, so that 80 percent of the terminal velocity is reached at 5 $R_\star$, i.e. 4 $R_\star$ above the surface. For increasing β, the velocity gradient in the wind becomes smaller. An alternative way of describing the velocity law is

$$v(r) = v_\infty\left(1 - \frac{\mathrm{r}_\circ}{r}\right)^\beta \tag{17.3}$$

where

$$\mathrm{r}_\circ = R_\star\left[1 - \left(\frac{v_\circ}{v_\infty}\right)^{1/\beta}\right] \tag{17.4}$$

This representation is often prefered as it is easier to handle in analytical considerations. Sometimes, for practical reasons, the velocity law is normalised to the terminal velocity, i.e.

$$w(x) \equiv v(x)/v_\infty \tag{17.5}$$

where the distance measure $x \equiv r/R_\star$ is in units of the stellar radius.

17.3 Flat topped and parabolic line shapes

We consider line formation in a stellar wind with an outflow velocity $v(r)$ that is monotonically increasing. An observer views the system along a line of sight that passes the center of the star at a distance p (see Figure 17.3); p is referred to as the impact parameter. The distance z on the impact beam is measured such that the positive z-direction is toward the observer and

$z = 0$ in the plane of the sky through the center of the star. The geometry is identical to that shown in Figure 3.6. Obvious relations are

$$r^2 = z^2 + p^2, \qquad \frac{z}{r} = \frac{v_z}{v} = \mu, \qquad \text{and} \qquad p = r\sqrt{1 - \left(\frac{v_z}{v}\right)^2}, \tag{17.6}$$

where $\mu = \cos\theta$ is the cosine of the angle between the radial direction and the beam direction (see also Figure 3.1).

Line-of-sight optical depth in a fast moving flow

The optical depth in the line, along this beam, is given by

$$\begin{aligned} \tau_{lu} &= \int_{-\infty}^{\infty} \alpha_{lu}(\nu)\, n_l\, dz \\ &= \left(\frac{\pi e^2}{m_e c}\right) f_{lu} \int_{-\infty}^{+\infty} n_l \left[1 - \frac{n_u\, g_l}{n_l\, g_u}\right] \phi(\Delta\nu)\, dz, \end{aligned} \tag{17.7}$$

where we have used Eq. (7.21) (more specifically Eq. 7.32) to express the cross section, and $\phi(\Delta\nu)$ is the line profile function. The profile function may, for instance, be given by Eq. (13.18). The frequency $\Delta\nu$ in the profile function depends on z through the Doppler relation (see Eq. 13.14)

$$\Delta\nu = \nu\left(1 - \frac{v_z}{c}\right) - \nu_{lu}. \tag{17.8}$$

We now assume that the geometrical zone along the z-axis in which the line is able to absorb the frequency ν is small. If the typical line width due to thermal broadening and turbulence is Δv, the extent of this zone is $\Delta z = \Delta v/(dv_z/dz)$. For Δz to be small, Δv should be small and the flow speed v should be large, such that dv_z/dz is large. If this zone is small, the level populations, source function, and line of sight velocity gradient (see below) will hardly vary and may be taken as constant. We refer to this zone as the *resonance zone* and may write

$$\tau_{lu} = \left(\frac{\pi e^2}{m_e c}\right) f_{lu}\, n_l \left[1 - \frac{n_u\, g_l}{n_l\, g_u}\right] \int_{-\infty}^{+\infty} \phi(\Delta\nu)\, dz \tag{17.9}$$

$$\tag{17.10}$$

The spatial integral can be reformulated in

$$\begin{aligned} \int_{-\infty}^{+\infty} \phi(\Delta\nu)\, dz &= \int_{+\infty}^{-\infty} \phi(\Delta\nu)\, \frac{\mathrm{d}z}{\mathrm{d}v_z}\, \frac{\mathrm{d}v_z}{\mathrm{d}\Delta\nu}\, d\Delta\nu = \frac{c}{\nu_{lu}}\, \frac{\mathrm{d}z}{\mathrm{d}v_z}\bigg|_r \int_{-\infty}^{+\infty} \phi(\Delta\nu)\, d\Delta\nu \\ &= \lambda_{lu}\, \frac{\mathrm{d}z}{\mathrm{d}v_z}\bigg|_r, \end{aligned} \tag{17.11}$$

where $dv_z/d\Delta\nu = -c/\nu_{lu}$ follows from inverting the derivative $d\Delta\nu/dv_z$ in Eq. (17.8) and taking $\nu = \nu_{lu}$. The last equality holds because the profile function is normalized to unity

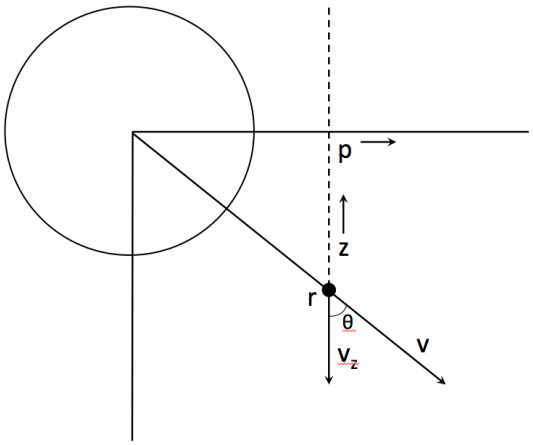


Figure 17.3: *Geometry in which the line flux is computed; identical to that shown in Figure 3.6.*

(see Eq. 7.4). Because we have assumed that in the resonance zone the line-of-sight velocity gradient dv_z/dz is constant, it may be taken out of the integral and evaluated at r.

For the optical depth, we then find

$$\tau_{lu} = \left(\frac{\pi e^2}{m_e c}\right) f_{lu}\, \lambda_{lu}\, n_l \left[1 - \frac{n_u\, g_l}{n_l\, g_u}\right] \left.\frac{\mathrm{d}z}{\mathrm{d}v_z}\right|_r = A(r) \left.\frac{\mathrm{d}z}{\mathrm{d}v_z}\right|_r . \tag{17.12}$$

$A(r)$ is a function of radius because n_l and n_u are. The velocity gradient dv_z/dz is given by

$$\begin{aligned} \frac{\mathrm{d}v_z}{\mathrm{d}z} &= \frac{\mathrm{d}}{\mathrm{d}z}\left(v\,\frac{z}{r}\right) = \frac{v}{r} + \frac{z}{r}\frac{\mathrm{d}v}{\mathrm{d}r}\frac{\mathrm{d}r}{\mathrm{d}z} - \frac{vz}{r^2}\frac{\mathrm{d}r}{\mathrm{d}z} \\ &= \frac{v}{r}\left[1 + \mu^2\left(\frac{r}{v}\frac{\mathrm{d}v}{\mathrm{d}r} - 1\right)\right] \\ &\equiv \frac{v}{r}\left[1 + \mu^2\sigma\right], \end{aligned} \tag{17.13}$$

where we have used that $dr/dz = z/r = \mu$. For a constant velocity stellar wind, we find that

$$\frac{\mathrm{d}v_z}{\mathrm{d}z} = \frac{v\,p^2}{r^3}. \tag{17.14}$$

An optically thin line in a constant velocity flow → flat-topped profile

To obtain the observed flux in the spectral line, we need to solve Eq. (3.30) for each line frequency. Note that an explicit reference to the frequency ν has disappeared from the expression for the optical depth and is replaced by a reference to the line-of-sight velocity v_z. This is typical for a moving medium. Think of an infinitely small resonance zone (we can do this easily

by assuming the intrinsic line profile function to be a delta function). In that case, a photon of frequency ν will be correspond to a unique line-of-sight velocity v_z (see Eq. 17.8 for the case $\Delta\nu = 0$). It is therefore the expression $\mathcal{F}_\nu(v_z)$ that will tell us how the flux changes throughout the line profile, i.e. what the line profile shape is.

Because the line source function in the resonance zone is constant, we have

$$\mathcal{F}_\nu = \frac{2\pi}{d^2}\int_0^R I_{v_z}(p)\,p\,dp = \frac{2\pi S_\nu}{d^2}\int_0^R \left(1-e^{-\tau_{lu}}\right)p\,dp, \tag{17.15}$$

where we have used Eq. (4.49) to write down the last equality and have assumed that the line source function S_ν is constant throughout the wind, i.e. $S_\nu \neq S_\nu(r)$. The medium beyond radius R does not contribute to the line flux. Note that strictly speaking, the above expression is not correct as we have assumed that none of the lines of sight hit the star. As long as the star is small compared to the volume of the wind envelope in which the line is formed, this will not be an issue. Instances in which this assumption is valid are ionic lines in Wolf-Rayet winds and molecular lines in the outflow of AGB stars and red supergiants. Because there is no background intensity for any of the beams, an emission line will form.

For a wind flow moving at constant velocity v, we may assume that (see Eq. 9.17)

$$A(r) = \frac{A_\circ}{v\,r^2}. \tag{17.16}$$

We then find for an optically thin line

$$\begin{aligned}
\mathcal{F}_\nu &= \frac{2\pi S_\nu}{d^2}\int_0^R \tau_{lu}\,p\,dp = \frac{2\pi S_\nu}{d^2}\int_0^R A\frac{r^3}{v\,p^2}\,p\,dp \\
&= \frac{2\pi\,A_\circ\,S_\nu}{v^2\,d^2}\int_0^R \frac{1}{r^2}\,\frac{r^3}{p}\,\frac{p}{r}\,dr = \frac{2\pi\,A_\circ\,S_\nu}{v^2\,d^2}\int_0^R dr \\
&= \frac{2\pi}{3}\frac{A_\circ\,R\,S_\nu}{v^2\,d^2},
\end{aligned} \tag{17.17}$$

where we have inserted result Eq. (17.14) and used that for constant v, $dp/dr = \sqrt{1-(v_z/v)^2} = \sqrt{1-(z/r)^2} = p/r$. The flux does not depend on v_z and therefore is constant for all line velocities $[-v,+v]$. The emission line is said to be block shaped or flat topped.

An optically thick line in a constant velocity flow → parabolic shaped profile

For an optically thick line in a flow moving at constant velocity, Eq. (17.15) reduces to

$$\begin{aligned}
\mathcal{F}_\nu &= \frac{2\pi S_\nu}{d^2}\int_0^R p\,dp = \frac{2\pi S_\nu}{d^2}\int_0^R \frac{p^2}{r}\,dr \\
&= \frac{2\pi S_\nu}{d^2}\int_0^R \left[1-\left(\frac{v_z}{v}\right)^2\right] r\,dr = \frac{2\pi S_\nu}{d^2}\left[1-\left(\frac{v_z}{v}\right)^2\right]\int_0^R r\,dr \\
&= \pi S_\nu\left(\frac{R}{d}\right)^2\left[1-\left(\frac{v_z}{v}\right)^2\right].
\end{aligned} \tag{17.18}$$

The emission line has a parabolic shape over the velocity range $[-v, +v]$.

17.4 The determination of mass loss from $\mathrm{H}\alpha$

We assume that we may consider the star and the stellar wind as two seperate entities. This is called the *core-halo approximation*. In reality, there obviously will be a fluent transition from photosphere to stellar wind. The star has a stellar radius $R_\star$ and emits a spectrum that is given by $\mathcal{F}_\nu(R_\star)$. We may expect the photospheric profile of $\mathrm{H}\alpha$ to be an absorption line. A spherically symmetric stellar outflow will either partly, fully, or more than fully fill in this this profile by $\mathrm{H}\alpha$ line emission. We assume that the $\mathrm{H}\alpha$ transition in the wind is optically thin such that all photons emitted in the spectral line will escape, save for those that happen to be emitted in the direction of the stellar disk. The latter photons will be absorbed by the star. Given the above assumptions the luminosity of the wind in $\mathrm{H}\alpha$ is given by (see eq. 7.16)

$$L(\mathrm{H}\alpha) = \int_{R_\star}^{\infty} [1 - W(r)]\ h\nu_{lu} n_u(r) A_{ul}\, 4\pi r^2\, dr \tag{17.19}$$

where u=3 and l=2 and ν_{lu} is 6563 Å. In this equation $W(r)$ is the geometrical dilution (eq 3.12), which corrects the volume integral for the fraction of total solid angle in which photons can not escape (as they will find the star on their way). The term $h\nu_{lu} n_u A_{ul}$ is the total energy that is emitted by the transition, in all directions and for all line frequencies, in erg cm^{-3} sec^{-1} (see § 7.3). The number density of particles in level n=3 can be written in terms of the electron density n_e and the proton density n_p, using Saha-Boltzmann equation (6.26) and the NLTE departure coefficient of the upper level b_u (see eq. 6.27). Manipulating a number of equations presented in § 6.5 we find (see eq. 6.32)

$$n_\mathrm{e} = \sum_k \sum_{j=1}^{J_k} j N_{jk} = N_N \sum_k A_k \sum_{j=1}^{J_K} j q_{jk} \equiv N_N \gamma \tag{17.20}$$

We further assume that the wind consists of a fully ionised hydrogen gas, such that $\gamma = 1$. The proton density is the product of the hydrogen ionisation fraction, the hydrogen number abundance and the number density of nuclei, i.e. $n_p = N_{11} = N_{jk} = q_{jk} A_k N_N$ (see § 6.5). In our case $n_p = N_N$. This results in

$$L(\mathrm{H}\alpha) = 4\pi\, h\nu_{lu} A_{ul} \int_{R_\star}^{\infty} [1 - W(r)]\ b_u\, \Phi_u(T) N_N^2\, r^2\, dr \tag{17.21}$$

where b_u is the departure coefficient of the upper level of the line (see eq. 6.26 and 6.27). The relation between the density of nuclei and the mass loss rate is described by eq. (9.17). If we also assume that the wind is isothermal and that the NLTE departure coefficient is constant throughout the wind, equation eq. (17.21) simplyfies to

$$L(\mathrm{H}\alpha) = \frac{h\nu_{lu} A_{ul}}{4\pi m_\mathrm{amu}^2}\, b_u \Phi_u(T) \frac{\dot{M}^2}{\mu_\mathrm{a}^2} \int_{R_\star}^{\infty} \frac{[1 - W(r)]}{r^2 v^2(r)}\, dr \tag{17.22}$$

In assuming a pure hydrogen gas we have fixed the mean atomic weight to $\mu_{\rm a} = 1$. If we normalize the velocity law to the terminal velocity (eq. 17.5) and switch to a dimensionless distance $x \equiv r/R_\star$, it follows, after substition of constants, that

$$L(\mathrm{H}\alpha) = 35.57\, b_u T^{-3/2} \exp(17538/T) \frac{\dot{M}^2}{R_\star v_\infty^2} \int_1^\infty \frac{[1 - W(x)]}{x^2 w^2(x)}\, dx \qquad L_\odot \qquad (17.23)$$

where $\dot{M}$ is in $M_\odot \mathrm{yr}^{-1}$; $R_\star$ in $R_\odot$, and v_∞ in $\mathrm{km\,sec}^{-1}$. The integral at the right side of this equation can easily be computed and is only a modest function of the free parameters $w_\circ \equiv v_\circ/v_\infty$ and β in the normalized velocity law (see eq. 17.2). For the wind temperature T we adopt a certain fraction of the effective temperature. Typically, it holds that $T/\,T_{\rm eff} \sim 0.5$ to 1. To be fair, a constant value for b_3 throughout the wind is a poor approximation, but, as we have chosen to go down this road, let us assume $b_3 \sim 1.3$. For O-type stars this is a fair compromise.

One can "measure" the quantity $L(\mathrm{H}\alpha)$ with relative ease: From the relations given in § 13.1 it follows that the luminosity in $\mathrm{H}\alpha$ produced by wind emission is given by

$$L(\mathrm{H}\alpha) = 4\pi d^2 \mathcal{F}^{\rm c}_{\mathrm{H}\alpha} \Big[W^{\rm phot}(\mathrm{H}\alpha) - W(\mathrm{H}\alpha) \Big] = 4\pi d^2 \mathcal{F}^{\rm c}_{\mathrm{H}\alpha} W^{\rm net} \qquad (17.24)$$

where $\mathcal{F}^{\rm c}_{\mathrm{H}\alpha}$ is the continuum flux at the wavelength 6563 Å; $W^{\rm phot}$ is the absorption equivalent with of the photospheric $\mathrm{H}\alpha$ profile, and W is the measured equivalent width. The net equivalent width, $W^{\rm net}$, i.e. the difference between the photospheric and measured equivalent width, provides a measure of the filling in of the spectral line by wind emission. If the effective temperature and gravity of the star are known we may obtain, for instance from Kurucz models, values for $\mathcal{F}^{\rm c}_{\mathrm{H}\alpha}$ and $W^{\rm phot}$. The stellar radius follows from the distance d. The terminal velocity may be determined from P Cygni profiles (see § 17.2).

Let us make a rough estimate of the photospheric flux in order to establish the way in which the net equivalent width depends on basic stellar parameters. For hot stars the $\mathrm{H}\alpha$ line is more or less in the Rayleigh-Jeans part of the spectrum. In that case we may use as a first order approximation of the photospheric flux

$$\mathcal{F}^{\rm c}_\lambda \sim 4\pi R_\star^2 \pi B_\lambda(\,T_{\rm eff}\,) \propto R_\star^2\, T_{\rm eff} \qquad (17.25)$$

If we assume the wind temperature to be a constant fraction of the effective temperature and that for hot O- and B-type stars the exponential term in eq. (17.23) is about unity, it follows that

$$W^{\rm net} \propto \left(\frac{\dot{M}}{R_\star^{3/2} v_\infty\, T_{\rm eff}^{5/4}} \right)^2 \qquad (17.26)$$

For an optically thin wind line the equivalent width is thus proportional to the square of the mass loss rate. The essence of the derived proportionality can easily be grasphed by realizing that the assumptions imply that $\mathrm{H}\alpha$ is formed by recombination processes, and that therefore the columns density will be proportional to $\sim \overline{\rho}^2 R_\star$. Applying continuity equation (9.14) the derived dependence is explained (save for the $T_{\rm eff}$ part).

17.5 The determination of mass loss from radio measurements

In 1975 Panagia & Felli and Wright & Barlow almost simultaneously pointed out that the mass loss from hot stars should also be measurable in infrared- and radio light. If a hot star has a stellar wind, it must be surrounded by an ionized plasma that can absorb and emit through bound-free and free-free transitions. The emission of the wind medium will cause the spectral energy distribution (SED) in the IR and radio to be flatter than the $\mathcal{F}_\nu \propto \nu^2$ one expects for a star without a wind (see eq. 6.9). In this section we concentrate on the radio as for this wavelength regime the observed flux can easily (read: analytically) be related to the mass loss.

For radio frequencies one may apply the Rayleigh-Jeans approximation ($h\nu/kT \ll 1$). Moreover, free-free processes are dominant. For a gas mixture the free-free extinction coefficient per cm is given by (see eq. 8.19)

$$\chi_\nu^{\rm ff} = 1.772 \times 10^{-2}\, \overline{g}_{\rm III}(\nu,T) \frac{Z^2\gamma}{T^{3/2}\nu^2}\, N_N^2 \tag{17.27}$$

$$\equiv K(\nu,T)\, N_N^2 \tag{17.28}$$

where we used eq. (17.20) to describe the electron density and

$$Z^2 \equiv \sum_k A_k \sum_{j=1}^{J_K} Z_{jk}^2 q_{jk} \tag{17.29}$$

for the quadratic mean of the charge of the ions. We also assume that the wind consists of a fully ionized hydrogen gas, such that $Z^2 = 1$ and $\gamma = 1$ (For this reason we have not added Z^2 en γ to the list of dependencies of K).

The optical depth along the line-of-sight at a projected distance $p > R_\star$ to the center of the stellar disk is (see § 3.4)

$$\tau_\nu^{\rm ff}(p) = K(\nu,T) \int_{-\infty}^{+\infty} N_N^2(p,z)\, dz \tag{17.30}$$

Using eq. (9.17) we get for the density of nuclei

$$N_N(r) = \frac{\dot{M}}{4\pi\, m_{\rm amu}\mu_{\rm a} r^2 v(r)} \simeq \frac{\dot{M}}{4\pi\, m_{\rm amu}\mu_{\rm a} r^2 v_\infty} \equiv \frac{A}{r^2} \tag{17.31}$$

For our hydrogen gas $\mu_{\rm a}$ = 1. In our case the almost equal sign in the above equation is justified because the bulk of the radio radiation will be emitted at large distances from the star. The optical depth further simplifies to

$$\tau_\nu^{\rm ff}(p) = K(\nu,T)A^2 \int_{-\infty}^{+\infty} \frac{dz}{r^4} = K(\nu,T)A^2 \int_{-\infty}^{+\infty} \frac{dz}{\left(p^2+z^2\right)^2} = K(\nu,T)A^2 \frac{c_2}{p^3} \tag{17.32}$$

In the last equality $c_2 = \pi/2$, which follows from

$$\int \frac{dz}{(p^2+z^2)^2} = \frac{1}{2}\frac{z}{p^2\,(p^2+z^2)} + \frac{1}{2}\frac{\arctan(z/p)}{p^3} \tag{17.33}$$

To obtain the emerging specific intensity along a beam specified by impact parameter $p > R_\star$ we assume that the wind is isothermal. In that case (see eq. 4.49)

$$I_\nu(p) = B_\nu(T)\left[1 - e^{-\tau_\nu^{\rm ff}(p)}\right] \tag{17.34}$$

The flux at distance d is given by (see eq. 3.30)

$$\begin{aligned}\mathcal{F}_\nu(d) &= \frac{2\pi}{d^2}\int_0^\infty I_\nu(p)p\,dp \\ &\simeq \frac{2\pi}{d^2}B_\nu(T)\left(K(\nu,T)A^2c_2\right)^{+2/3}\int_0^\infty \left[1-e^{-1/y^3}\right] y\,dy \end{aligned}\tag{17.35}$$

where

$$y \equiv p\left(K(\nu,T)A^2c_2\right)^{-1/3} \tag{17.36}$$

The the last term in eq. (17.35) is not identical to the previous term because the optical depth along lines of sight intercepted by the star, i.e. for $p \le R_\star$, are strictly speaking, not given by eq. (17.32) (in this case the integration should run from $\sqrt{R_\star^2 - p^2}$ to $+\infty$), and $I_\nu(p)$ should not be represented by eq. (17.34); see eq. 4.49). However, as the radio radiation is produced in a large volume around the star this introduces a negligible error. The y-integral can be solved numerically, from which results $\mathcal{C}_2 = 1.339$. For the radio regime ($h\nu/kT \ll 1$) we obtain

$$\mathcal{F}_\nu(d) = 23.454\,\frac{\nu^{2/3}}{(d/\,\mathrm{kpc})^2}\left(\frac{\overline{g}_{\rm III}(\nu,T)\gamma Z^2}{\mu_{\rm a}^2}\right)^{2/3}\left(\frac{\dot{M}/\,M_\odot\mathrm{yr}^{-1}}{v_\infty/\,\mathrm{km\,sec}^{-1}}\right)^{4/3} \quad \mathrm{Jy} \tag{17.37}$$

The flux is given in jansky. Notice that in our isothermal wind $\mathcal{F}_\nu$, save for a weak dependence which enters through the Gaunt factor, is independent of the temperature of the gas. This is important in view of uncertainties in the temperature structure of stellar winds. If we also neglect the modest frequency dependence of $\overline{g}_{\rm III}$ we find for *spectral slope* or *spectral index* $n \equiv \partial\log\mathcal{F}_\nu/\partial\log\nu = 2/3$, i.e.

$$\mathcal{F}_\nu \propto \nu^{2/3} \tag{17.38}$$

A good approximation for the Gaunt factor in the radio domain is provided by

$$\overline{g}_{\rm III} = 9.77\,\left(1 + 0.13\log[T^{3/2}/\nu Z]\right) \tag{17.39}$$

Corrected for this behaviour we obtain $\mathcal{F}_\nu \propto \nu^{0.6}$. In the infrared $\overline{g}_{\rm III}$ is almost independent of frequency and the spectral slope remains as implied by eq. (17.38). Inversion of eq. (17.37) yields for the mass loss

$$\dot{M} = 0.094\,\frac{(d/\,\mathrm{kpc})^{3/2}}{\nu^{1/2}}\,\frac{\mu_{\rm a}}{(\gamma\overline{g}_{\rm III})^{1/2}Z}\left(v_\infty/\,\mathrm{km\,sec}^{-1}\right)\left(\mathcal{F}_\nu/\mathrm{Jy}\right)^{3/4} \quad M_\odot\mathrm{yr}^{-1} \tag{17.40}$$

The insensitivity of the radio flux for the temperature of the wind and the stellar radius make that mass loss rates derived from radio measurements are among the most reliable. Unfortunately, the number of stars for which we can use the radio method is small as the fluxes are low, therefore difficult to measure. For a star at a distance of 1 kpc, having $\dot{M} = 10^{-6}\ M_{\odot}\mathrm{yr}^{-1}$ and $v_{\infty} = 2000\ \mathrm{km\,sec}^{-1}$, the flux at 6 cm is $\mathcal{F}_{\nu} \simeq 0.3$ mJy. The detection limit of large radio telescopes is near 0.1 mJy. The radio method can therefore only be applied for relatively nearby stars (d $\lesssim$ 1 kpc) with relatively high mass loss ($\dot{M} \gtrsim 10^{-6}\ M_{\odot}\mathrm{yr}^{-1}$).

Effective radius at radio wavelengths

Based on the measured ratio $\dot{M}/v_{\infty}$ we can assign to a star an effective radius at radio wavelengths. The value for the radius is a function of wavelength at which it is measured and depends on the exact definition of r_{ν}^{ff}.

The optical depth at the effective radius r_{ν}^{ff} is (compare to eq. 17.32)

$$\tau_{\nu}^{\mathrm{ff}}(r_{\nu}^{\mathrm{ff}}) = K(\nu,T)A^2 \int_{r_{\nu}^{\mathrm{ff}}}^{\infty} \frac{dr}{r^4} = \frac{KA^2}{3(r_{\nu}^{\mathrm{ff}})^3} \tag{17.41}$$

We discuss two possible definitions of the effective radius. In the first case we focus our attention to the medium above the radio surface of the star and assume that this radiating medium is optically thin; in the second case we require the emission to come from an optically thick radiating sphere.

If we require the observed flux to be equal to the total emission in the wind above a certain effective radius r_{ν}^{ff}, i.e.

$$L_{\nu} = 4\pi d^2\, \mathcal{F}_{\nu}(d) = 4\pi \left(r_{\nu}^{\mathrm{ff}}\right)^2\, \mathcal{F}_{\nu}(r_{\nu}^{\mathrm{ff}}) = \int_{r_{\nu}^{\mathrm{ff}}}^{\infty} 4\pi\eta_{\nu}^{\mathrm{ff}}\, 4\pi r^2\, dr \tag{17.42}$$

we find using $\eta_{\nu}^{\mathrm{ff}} = \chi_{\nu}^{\mathrm{ff}} B_{\nu}(T)$, substitution of eq. (17.41) and equating with eq. (17.35), that the effective radius should be measured at $\tau_{\nu}^{\mathrm{ff}} = 0.247$. Notice that in this definition we have implicitely assumed that the radiating medium is optically thin. Substitution of constants and inversion of eq. (17.41) leads to

$$r_{\nu}^{\mathrm{ff}} = 3.633 \times 10^{17} \frac{1}{(\tau_{\nu}^{\mathrm{ff}})^{1/3}} \frac{1}{\nu^{2/3}T^{1/2}} \left(\frac{\overline{g}_{\mathrm{III}}(\nu,T)\gamma Z^2}{\mu_{\mathrm{a}}^2}\right)^{1/3} \left(\frac{\dot{M}/\,M_{\odot}\mathrm{yr}^{-1}}{v_{\infty}/\,\mathrm{km\,sec}^{-1}}\right)^{2/3} R_{\odot} \tag{17.43}$$

If we again take $\dot{M} = 10^{-6}\ M_{\odot}\mathrm{yr}^{-1}$ en $v_{\infty} = 2000\ \mathrm{km\,sec}^{-1}$ and assume that the temperature of the wind $T = 30\,000$ K we find at the wavelength of 6 cm an effective radius $r_{\nu}^{\mathrm{ff}} = 1308 R_{\odot}$. The typical stellar radius of an O star (measured at optical wavelengths) is about 15 $R_{\odot}$ (see table B.5) which shows that the radio radius $r_{\nu}^{\mathrm{ff}} \gg R_{\star}$.

We may also define the effective radius at radio wavelengths by equating the measured radio flux with the total amount of radiation emitted by a spherical surface of radius r_{ν}^{ff}, i.e. (see

vgl. 3.27)

$$\mathcal{F}_\nu(d) = \pi B_\nu(T) \left(\frac{r_\nu^{\rm ff}}{d} \right)^2 \tag{17.44}$$

After substitution of eq. (17.41) and equating to eq. (17.35) it follows that the effective radius should be measured at $\tau_\nu^{\rm ff} = 0.0484$. Note that in this definition the emission is thought to originate from an optically thick surface. The reason why the "optically thick" effective radius is larger than the "optically thin" one has to do with the fact that the latter requires an integration over a large volume.

Exercise 17.1

Measure the terminal velocity of the wind of ζ Puppis, using the spectrum displayed in Fig. 17.2, from the saturated C IV $\lambda\lambda$1448,1551 doublet resonance line. Use the pdf version of these lecture notes to zoom in on the profile to enlarge the region around the C IV profile to improve the accuracy of your measurement.

Exercise 17.2

Sketch the profile of an inflow sensitive line, formed through resonance scattering, of a star that is accreting material through spherical infall. In this process, the gas accelerates from a negligibly small infall velocity $v \sim 0$ far away from the star to a maximum infall velocity v_∞ at the stellar surface. Assume that the radius of the star is much smaller than the volume surrounding the star from which the emission originates.

Exercise 17.3

Let us consider a spectral line that is formed in an extended stellar wind. The line formation volume is so large that the star may be considered a point source. The line is formed through pure scattering and shows a P Cygni profile.

a) Is the emission equivalent width of this line (i.e. the part of the profile surface above the stellar continuum larger/equal/smaller to the absorption equivalent width (i.e. the part of the profile surface that is below the stellar continuum? Explain your choice.

Let us now consider a case where the line formation volume is small compared to the dimension of the star.

b) Is the emission equivalent width larger/equal/smaller than the absorption equivalent width? Explain your choice.

Exercise 17.4

This could be a nice exam question. In this exercise we aim to determine the velocity law for a stationairy, i.e. time-independent, radial stellar wind that is driven *solely* by the force due to the gradient of the radiation pressure (see eq. 9.35). We assume that the extinction can be expressed using the flux-weighted mean extinction (see eq. 10.21). This implies that we focus on continuum driven winds as it is not allowed to express the

radiation pressure on spectral lines using a grey extinction. At the stellar radius, the flow has velocity $v(R_\star) = v_\circ$. The equation of motion is given by Eq. 9.26, i.e.,

$$\rho \frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} = \rho \left[\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla)\mathbf{v} \right] = -\nabla p + \mathbf{f}, \tag{17.45}$$

where d/dt is the fluid-frame time derivative (see Eqs. 9.14 and 9.24) and $\partial/\partial t$ the partial time derivative at a fixed point in space. For a time-independant radial flow this implies

$$\frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} = \frac{\partial v}{\partial t} + v \frac{\partial v}{\partial r} = v \frac{\mathrm{d}v}{\mathrm{d}r}, \tag{17.46}$$

where we have switched to total derivative notation (d/dt) in the last equality as changes in velocity are only traced in one direction.

a) Give an expression for the ratio, Γ, between the acceleration due to radiation pressure and gravity

b) Show that the equation of motion of our problem is given by

$$v \frac{\mathrm{d}v}{\mathrm{d}r} = -\frac{GM_\star(1-\Gamma)}{r^2} \tag{17.47}$$

c) Solve the equation of motion.

d) Take $v_\circ = 0$. (Why is this not physical?). Give the solution $v(r)$ in terms of the escape velocity, v_{esc}, at the stellar radius. Which value of β describes the steepness of the velocity law?

e) Give an expression for the terminal flow velocity in terms of Γ and v_{esc}. Which values of Γ assure a physical solution?

Exercise 17.5

Assume a geometrically thin, homogeneous shell of radial thickness Δr moving at constant velocity v at distance $r \gg R_\star$ from the central star having radius $R_\star$. This is so far from the star that for practical purposes we may consider the star to be a point source. What will be the profile shape of an emission line formed in this shell?

Exercise 17.6

Assume a geometrically and optically thin shell at distance $r \gg R_\star$ from a central star having radius $R_\star$, moving at radial velocity v. Show that emission from all redshifted velocities $-v$ to $-v\left[1 - 1/2\left(R_\star/r\right)^2\right]$ is obscured by the stellar disk.

18

H II regions

In this chapter we move away from the stellar surface into the ambient interstellar medium of O- and early B-type stars. As we have seen in chapter 2 these massive stars are both extremely luminous ($L_\star \gtrsim 10^4 L_\odot$) and extremeley hot, with surface temperatures $T_{\rm eff} \gtrsim 20\,000$ K. The extreme ultraviolet radiation that is the result of these stellar properties causes ionization of the surrounding gas. As the ionized gas is predominantly hydrogen these regions are known as H II regions. The balance between photo-ioniziation and radiative recombination determines the degree of ionization. The kinetic energy of the photo-electrons (those kicked out of the atom or ion during ionization) is quickly shared with other free electrons in the nebula, establishing a Maxwellian energy distribution. Electrons from the thermal pool can excite low-lying levels of metals, such as O III, and downward radiative transitions cool the nebula. This energy balance sets the temperature of the gas.

The processes setting the ionization and temperature structure of nebulae represent extreme NLTE conditions; studying these conditions provides important insight in NLTE physics and for this region this topic is part of these lecture notes. We first give a brief historical introduction. This is followed by a discussion of what sets the size of an H II-region. We continue with a discussion on hydrogen line radiation and hydrogen free-free radio continuum emission. Finally, we discuss emission from metal lines.

18.1 Historical introduction

In the 19th century an animated debate ensued on the question whether the diffuse objects seen through telescopes, loosely termed 'nebulae', were in fact unresolved stellar associations or gas clouds. The answer came when these objects were studied using spectroscopy. The small 'white nebulae', those located primarily outside of the plane of the Milky Way, featured spectra typical for that of stars. These are stellar associations or star clusters. Other nebulae showed a pronounced line spectrum, sometimes accompanied by an underlying continuum that did not appear to be of stellar origin. These are the gas clouds, including planetary nebulae, H II regions, and supernova remnants.

Figure 18.1: *The Orion Nebula (also known as Messier 42, M42, or NGC 1976) at a distance of ∼400 pc is the closest region of massive star formation to Earth. The dimensions of this image are* 65×60 *arcmins, corresponding to a physical scale of* 7.5×7.0 *pc. The open cluster of stars in the heart of the Orion Nebula contains several O-type stars of which the O6pe V star* θ^1 *Orionis C is the most massive. The optical light we see is the result from line emission of hydrogen (Balmer lines) and reflection of star light on dust grains.*

The spectrum of the brightest of these nebulae, the Orion nebula (see Fig. 18.1), was observed in 1863 by William Huggins. Soon after, Balmer lines were recognized in this and other nebulae. After it was discovered in the Sun, also helium was found. Surprisingly enough, the identification of the by far strongest lines in many nebular spectra – in green light at 4959 and 5007 Å – and those of other strong transitions, turned out to be an enormous challenge. Some even speculated that these unidentified lines were produced by a new and hypothetical element 'nebulium'. Progress in the understanding of atomic structure eventually led to the identification of these lines as forbidden transitions of O III. Other strong nebular lines were found to be of a similar nature, and due to known elements such as nitrogen, neon, sulfer and argon. The complexity of the spectrum of H II regions is nicely illustrated in Fig. 18.2, that

Tielens, page 252

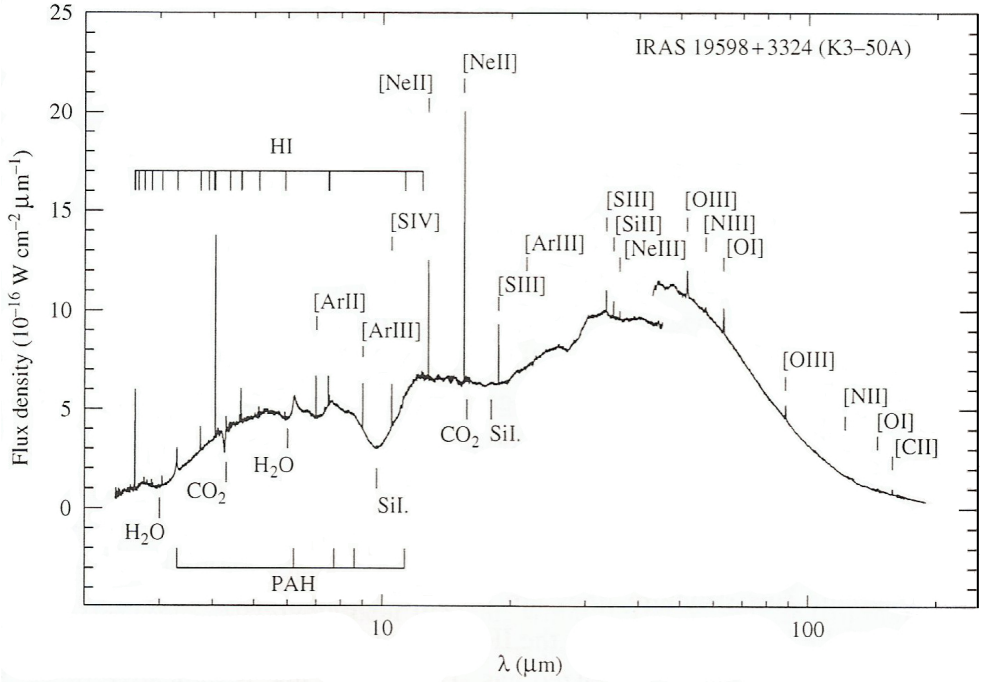


Figure 18.2: *The infrared spectrum of the H II region K3-50, measured by the Infrared Space Observatory. This spectrum shows a multitude of components. The continuum emission is due to warm dust. Superimposed, are IR emission bands due to PAH molecules. Absorption bands due to silicates and ices located in a foreground cold molecular cloud are also visible. The line spectrum shows hydrogen recombination lines and far-IR ionic fine structure lines originating in the H II regions, as well as those originating in the photo-dissociation region. The jump in the spectrum reflects the difference is aperture size between ISO's short wavelength and long wavelength spectrometer and is not intrinsic. Adapted from: Peeters et al. (2002).*

shows the IR spectrum of K3-50.

H II regions are formed when massive stars reach the main sequence and suddenly 'switch on' their UV radiation field. The gas and dust that remains from the star formation process is often still around and will be ionized by the UV radiation. Regions of ionized gas around O and B stars are therefore often called H II regions. *Planetary nebulae* have many things in common with H II regions: there too a gas/dust mixture (the stellar wind which was ejected when the star was still an asymptotic giant branch star) is ionized by the UV photons of the hot white dwarf. White dwarfs can reach much higher temperatures than main sequence stars, so in general one can expect higher ionization of the gas in planetary nebulae.

18.2 The primary radiation mechanism in nebulae

The first thing to realize when identifying the relevant physical processes in a gaseous nebula is the large distance of the nebula to the hot star (or stars) that irradiates (or irradiate) the gas. This causes the radiation field, though very 'hot', to be very diluted.

Atomic processes that depend on radiation, such as photoexcitation and photoionization, thus proceed at a rate that is a factor W (see eq. 3.12, or for a more practical version eq. 15.3) slower than in thermodynamic equilibrium. However, the rate of recombinations is independent of W and is controlled by local properties only, that is, the electron and ion densities and the temperature of the medium (see vgl. 8.8). To give an example, a hydrogen atom in an excited state will almost certainly experience a spontaneous de-excitation to a lower level before it can absorb a photon that could bring it in a higher excited state, or cause ionization. The result is that almost all the neutral hydrogen will be in the ground state. All H ionizations by photons having $\lambda < 912$ Å will thus arise from the ground level, whilst recombinations may happen to any level. When the capture of a free electron leads to a recombination to the ground level, a Lyman continuum photon will be emitted. When recombination to $n = 2$ takes place, a Balmer continuum photon will be released ($\lambda < 3646$ Å), followed by a Lyα photon when the atom returns to the ground state. An electron captured in level $n = 3$ will discharge a Paschen continuum photon ($\lambda < 8204$ Å). There are now two possibilities. Either the atom de-excites to $n = 2$ by emitting a Hα photon and subsequently to $n = 1$ by releasing a Lyα photon, or it cascades directly to the ground level by emitting a Lyβ photon. Scenarios for recombinations to $n = 4$ or higher levels can easily be worked out (see figure 6.4).

18.3 Ionization equilibrium

The ionization equilibrium of the elements is determined by the balance between photoionizations and recombinations of electrons with ions. Let us assume that the nebula is composed of hydrogen only. What then is the ionization condition of H? Given the strong dilution of the radiation field (see above) and the large excitation energies of hydrogen, it is safe to assume that all the neutral hydrogen is in the ground state, i.e. $n_0 = N^0$. Using Eq. (8.1), (8.3) and (8.8), we may write (see Eq. 15.4)

$$n_0\, 4\pi \int_{\nu_\circ}^{\infty} \alpha_\nu^{\rm bf} \frac{J_\nu}{h\nu}\, d\nu = N^+ n_{\rm e}\, \alpha_{\rm A}(T) \tag{18.1}$$

where $\nu_\circ$ corresponds to 912 Å, and $\alpha_{\rm A}$ denotes the total recombination coefficient of hydrogen to all levels (see Table 18.1). If $q^0 = N^0/N$ is the fraction of neutral hydrogen, it then follows that $N^+ = n_{\rm e} = (1 - q)$. We further assume that the radiation field in the nebula is completely controlled by a star at a large distance. If we adopt the star to be an isotropically

radiating sphere, it follow from Eq. (3.11) and (3.27) that

$$J_\nu(r) = \frac{1}{4\pi}\mathcal{F}_\nu(r) = \frac{1}{4\pi}\frac{L_\nu}{4\pi r^2} \tag{18.2}$$

Writing for the total number of Lyman continuum photons leaving the star each second

$$Q_0 \equiv \int_{\nu_\circ}^{\infty} \frac{L_\nu}{h\nu}\, d\nu \tag{18.3}$$

and approximating the extinction coefficient of photoionization with a constant (frequency independent) value $\overline{\alpha}_\nu^{\rm bf}$, we obtain for the fraction of neutral hydrogen

$$\frac{q}{(1-q)^2} = \frac{4\pi r^2 \alpha_{\rm A} N}{\overline{\alpha}_\nu^{\rm bf} Q_0}. \tag{18.4}$$

Let us consider a point in the nebula with a typical particle density $N = 10\,{\rm cm}^{-3}$, at a distance of 5 pc from an O7 V star having $\log Q_0 = 48.7$. We then find – using $\overline{\alpha}_\nu^{\rm bf} \sim 6 \times 10^{-18}$ (see Eq. 8.3) and $\alpha_{\rm A} \sim 4 \times 10^{-13}$ (see Table 18.1) – for the fraction of neutral hydrogen $q \sim 4 \times 10^{-4}$. In other words: hydrogen is almost completely ionized.

The recombination time of nebular gas

The recombination time of the nebular gas is given by

$$t_{\rm rec} = \frac{1}{n_{\rm e}\, \alpha_{\rm A}} \tag{18.5}$$

For $n_{\rm e} = 10\ {\rm cm}^{-3}$ (assuming near complete hydrogen ionization) the recombination time is $\sim 3 \times 10^{11}$ sec or $\sim 10^4$ yr for a typical nebular temperature (see Table. 18.1). Thus, in such a region, if a neutral hydrogen atom is photoionized, it stays ionized for a long time before recombining. Once it recombines, it is relatively quickly ionized again, and so the instantaneous neutral fraction in this region should be very small.

18.4 Strömgren sphere

A source that is producing a finite number of ionizing photons can of course not ionize an infinitely large volume. Therefore, if the star is located in a sufficiently large gas cloud, there must be a boundary to the H II region. Somewhere there will be a zone in which hydrogen is recombining and outside of which there is an H I region. The thickness of the transition zone will be small because when hydrogen starts to recombine the optical depth in the Lyman continuum rapidly increases, effectively preventing ionizing radiation to penetrate deeper into the medium. Hence we may estimate the thickness of the recombination zone by equating it

Table 18.1: *Recombination coefficients of hydrogen as a function of temperature in cm^3 sec^{-1}: $\alpha_{\rm A} \equiv \sum_{i=0}^{\infty} \alpha_{i,H}^{RR}$ gives the total number of recombinations to all levels; $\alpha_{\rm B} \equiv \sum_{i=1}^{\infty} \alpha_{i,H}^{RR}$ gives the total number of recombinations to all levels above the ground level. From: Storey & Hummer (1995), for $n_{\rm e} = 1000$. The fit formula is from Draine (2011), in which $T_4 = T/10\,000$ and Z the net charge of the ion. It provides good approximations in the range $30\,{\rm K} \leq T/Z^2 \leq 30\,000\,{\rm K}$.*

T	$\alpha_{\rm A}/\,10^{-13}$	$\alpha_{\rm B}/\,10^{-13}$
3 000	9.74	6.74
5 000	6.83	4.53
10 000	4.17	2.59
15 000	3.11	1.84
20 000	2.51	1.43
30 000	1.84	0.991
T	$\sim 4.15\,(T/10\,000\,{\rm K})^{-0.72}$	$\sim 2.60\,(T/10\,000\,{\rm K})^{-0.80}$

to the mean free path at the point where half of the hydrogen is recombined, i.e. at $q = 0.5$. We find (see Eq. 4.28) using the numbers that we have used above

$$\Delta r \sim \frac{1}{\overline{\alpha}_\nu^{\rm bf}\, N^0} = \frac{1}{\overline{\alpha}_\nu^{\rm bf}\, q\, N} \sim 0.01\,{\rm pc}. \tag{18.6}$$

This is much smaller than the typical radius of an H II region. H II regions thus have a sharp boundary, referred to as the *ionization front*. Inside of the boundary hydrogen is almost completely ionized; outside of the boundary hydrogen is almost completely neutral.

A source that is producing a finite number of ionizing photons can of course not ionize an infinitely large volume. Therefore, if the star is located in a sufficiently large gas cloud, there must be a boundary to the H II region. Somewhere there will be a zone in which hydrogen is recombining and outside of which there is an H I region. The thickness of the transition zone will be small because when hydrogen starts to recombine the optical depth in the Lyman continuum rapidly increases, effectively preventing ionizing radiation to penetrate deeper into the medium. Hence we may estimate the thickness of the recombination zone by equating it to the mean free path at the point where half of the hydrogen is recombined, i.e. at $q = 0.5$. We find (see Eq. 4.28) using the numbers we have used above

$$\Delta r \sim \frac{1}{\overline{\alpha}_\nu^{\rm bf}\, N^0} = \frac{1}{\overline{\alpha}_\nu^{\rm bf}\, q\, N} \sim 0.01\,{\rm pc}. \tag{18.7}$$

This is much smaller than the typical radius of an H II region (see below). H II regions thus have a sharp boundary, referred to as the *ionization front*. Inside of the boundary hydrogen is almost completely ionized; outside of the boundary hydrogen is almost completely neutral.

Strömgren sphere – optically thin Lyman continuum

How big is an H II region? Let us assume a homogeneous spherical cloud of hydrogen gas with a star in the center, and that the cloud is transparent for Lyman continuum photons. In very good approximation, the sphere up to the ionization front is almost fully ionized. We can thus write $N^+ \simeq n_\mathrm{e} \simeq N$. If R_S is the radius of the H II region, the balance between the number of ionizations per second and the number of recombinations per second is

$$Q_0 = \frac{4\pi}{3} R_\mathrm{S}^3 \, \alpha_\mathrm{A} \, n_\mathrm{e} \, N^+ = \frac{4\pi}{3} R_\mathrm{S}^3 \, \alpha_\mathrm{A} \, N^2. \tag{18.8}$$

Hence, the radius of the sphere is

$$R_\mathrm{S} = \left(\frac{3}{4\pi}\right)^{1/3} \left(\frac{Q_0}{\alpha_\mathrm{A}}\right)^{1/3} N^{-2/3}. \tag{18.9}$$

This radius is referred to as the *Strömgren radius*, in honor of the pioneer of this field, and the volume encompassed by R_S as the *Strömgren sphere*. It shows that the higher the gas density, the smaller is the N II region.

If we fill in the numbers of our standard example (see above), we find $R_\mathrm{S} = 10.1$ pc. Indeed, the Strömgren radius is much larger than the thickness of the ionization front.

Strömgren sphere – optically thick Lyman continuum

In the previous discussion we assumed the Strömgren sphere to be optically thin and the radiation field in the nebula to be completely controlled by the central star. However, at Lyman continuum frequencies $\nu \geq \nu_\circ$ the nebula very likely is not optically thin. Therefore, we now allow the mean intensity to have a diffuse component as a result of Lyman continuum radiation emitted in the nebula. The local volume emission coefficient η_ν for radiation generated by recombinations to the ground state is given by

$$4\pi \int_{\nu_\circ}^{\infty} \frac{\eta_\nu}{h\nu} \, d\nu = N^+ \, n_\mathrm{e} \; \alpha_{1,H}^{RR}, \tag{18.10}$$

where $\alpha_{1,H}^{RR}$ is the recombination coefficient to the ground level of hydrogen.

Suppose the nebula is optically thick for Lyman radiation, such that no Lyman continuum photons can escape. Each and every Lyman photon produced by the diffuse radiation field will be absorbed elsewhere in the nebula. For the mean intensity of the diffuse component of the radiation field, J_ν^d, this implies

$$n_0 \, 4\pi \int_\mathrm{nebula} \alpha_\nu^\mathrm{bf} \, \frac{J_\nu^\mathrm{d}}{h\nu} \, dV = 4\pi \int_\mathrm{nebula} \frac{\eta_\nu}{h\nu} \, dV \tag{18.11}$$

where the integration is over the entire volume of the nebula. If the mean free path of the photons is small, that is, if the photons are absorbed close to where they are created, one may

even write

$$J_\nu^{\rm d} = \frac{\eta_\nu}{n_0\,\alpha_\nu^{\rm bf}}. \tag{18.12}$$

This is referred to as the *on-the-spot approximation*. For nebulae having particles densities that are not too small this is a reasonable assumption.

As a result of absorptions (read: photo-ionizations) Lyman continuum radiation that originates directly from the star gets weakened by a factor $\exp[-\tau_\nu(r)]$, where τ_ν is the radial optical depth

$$\tau_\nu(r) = \int_0^r n_0\,\alpha_\nu^{\rm bf}\,dr. \tag{18.13}$$

In a formal sense, the integration of course needs to start at $R_\star$ in stead of at the origin. However, as we will show below, things work out more elegantly if we accept this minute error. For the mean intensity of the stellar component of the radiation field we may write (using also Eq. 18.2)

$$J_\nu^{\rm s}(r) = \frac{1}{4\pi}\,\frac{L_\nu}{4\pi r^2}\,e^{-\tau_\nu}. \tag{18.14}$$

The total radiation field in the nebula is then the sum of the stellar and the diffuse component, i.e. $J_\nu = J_\nu^{\rm s} + J_\nu^{\rm d}$. Substituting Eq. (18.12) and (18.14) in ionization equilibrium (18.1) then yields

$$n_0 \int_{\nu_\circ}^\infty \alpha_\nu^{\rm bf}\,\frac{L_\nu}{h\nu}\,\frac{e^{-\tau_\nu}}{4\pi r^2}\,d\nu = N^+\,n_{\rm e}\left[\alpha_{\rm A} - \alpha_{1,H}^{RR}\right] = N^+\,n_{\rm e}\,\alpha_{\rm B}, \tag{18.15}$$

where we have used Eq. (18.10). We have introduced the recombination coefficient $\alpha_{\rm B} = \sum_2^\infty \alpha_{i,{\rm H}}^{\rm RR}$ to represent the total number of recombinations to all levels above the ground level (see Table 18.1). The physical meaning of the above equation is that in an optically thick nebula ionizations caused by the stellar radiation field are in equilibrium with recombinations to excited levels of hydrogen. Recombinations to the ground level produce Lyman continuum photons that (in the on-the-spot approximation) are almost on-the-spot reabsorbed and thus have no effect on the ionization equilibrium.

To determine the radius of the ionized region we use that $d\tau_\nu = n_0\,\alpha_\nu^{\rm bf}\,dr$ (see Eq. 18.13) and integrate over distance. This yields

$$\int_{\nu_\circ}^\infty \frac{L_\nu}{h\nu}\left[\int_0^\infty e^{-\tau_\nu}\,d\tau_\nu\right]\,d\nu = \int_{\nu_\circ}^\infty \frac{L_\nu}{h\nu}\,d\nu = \int_0^\infty N^+\,n_{\rm e}\,\alpha_{\rm B}\,4\pi r^2\,dr \tag{18.16}$$

By letting the optical depth run from zero to infinity we have assured that all Lyman continuum photons are 'used up'. To find a simple expression for the size of the ionized region we again assume that the gas is fully ionized up to the Strömgren radius $R_{\rm S}$, and that the gas outside of this region is neutral. As we have seen, we expect a classical H II region to have a sharp boundary implying that our assumption is quite reasonable. Using Eq. (18.3) it follows that

$$Q_0 = \frac{4\pi}{3} R_{\rm S}^3\,\alpha_{\rm B}\,N^2 \tag{18.17}$$

The meaning of this result is that the total number of ionizing photons emitted by the star (per second) is exactly balanced by the total number of recombinations to excited states of all hydrogen atom within a Strömgren sphere. For the Strömgren sphere we find

$$R_{\rm S} = \left(\frac{3}{4\pi}\right)^{1/3} \left(\frac{Q_0}{\alpha_{\rm B}}\right)^{1/3} N^{-2/3} = 1.6\,{\rm pc} \left(\frac{Q}{10^{50}\,{\rm s}^{-1}}\right)^{1/3} \left(\frac{N}{10^3\,{\rm cm}^3}\right)^{-2/3} \tag{18.18}$$

If we again fill in the numbers of our standard example (see above) and adopt $\alpha_{\rm B} = 2 \times 10^{-13}$ cm^3 s^{-1}, we find $R_{\rm S} = 12.7$ pc.

The two assumptions that we have made to estimate the size $R_{\rm S}$ of the H II region are the on-the-spot approximation and that the medium is either fully ionized (within the Strömgren sphere) or fully neutral (outside of the Strömgren sphere). In doing so we could avert having to solve the equation of transfer explicitly.

Ionization bounded and density bounded

Note that we have assumed that the ionizing star is in a sufficiently large gas cloud, that is, that the edge of the Strömgren sphere is determined by the region where hydrogen recombines. This is referred to as *ionization bounded*. It is however possible that the nebula contains insufficient hydrogen for all Lyman continuum photons to 'be used'. In that case we refer to the H II region as *density bounded*.

What is the total mass of ionized hydrogen gas in an ionization bounded nebula. This is

$$M = \frac{4\pi}{3} R_{\rm S}^3\, m_{\rm amu}\, N = \frac{Q_0\, m_{\rm amu}}{\alpha_{\rm B}\, N} \simeq 417\, M_\odot \left(\frac{Q}{10^{50}\,{\rm s}^{-1}}\right) \left(\frac{10^3\,{\rm cm}^{-3}}{N}\right), \tag{18.19}$$

where we have adopted Case B. For our standard example we find $M = 2089\ M_\odot$.

Case A and Case B recombination

While we are in the process of defining things: in case the nebula is optically thin in all recombination lines, i.e. all radiation produced by recombination processes in the nebula is able to escape freely, one speaks of *Case A*. H II regions that fulfill the Case A requirement can only contain a relatively small amount of gas. So little gas, as a matter of fact, that they will be hard to observe.

Nebulae containing appreciable amounts of gas will rapidly develop large optical depths in the hydrogen Lyman lines. We can easily estimate this from the ratio of the extinction coefficients per particle in the ground state for line- and continuum radiation. For the extinction coefficient of lines we have

$$\chi_\nu = \alpha_{lu}(\nu)\, n_l = \frac{\pi e^2}{m_{\rm e} c}\, f_{lu}\, \phi_\nu\, n_l = \frac{h\nu}{4\pi}\, B_{lu}\, \phi_\nu\, n_l, \tag{18.20}$$

where $(\pi e^2/m_e c)$ = 0.02654 cm^2 hz and f_{lu} is the oscillator strength (see Sect. 7.3). Comparison of the collisional cross section at line center $\alpha_{lu}(\nu_\circ)$, assuming a Doppler function for the profile function ϕ_ν, with the continuum cross section Eq.(8.3) at the ionization boundary yields, after substitution of constants

$$\frac{\tau(\mathrm{Ly}\,n)}{\tau(\mathrm{Ly{-}edge})} = 14.68\,\frac{\lambda_{1n}\,[\mathring{A}]\,f_{1n}}{g_{\mathrm{II}}(\nu_\circ, n)\,(T/10\,000)^{1/2}}, \tag{18.21}$$

where n is the principle quantum number of the upper level of the line that is considered. For a characteristic temperature $T = 10\,000$ K and $g_{\mathrm{II}} \sim 1$ we find that Lyα has about a 10^4 times larger optical depth than the continuum at the Lyman ionization boundary. For an ionization bounded nebula with τ(Ly-edge) ~ 1 one thus obtains τ(Lyα) $\sim 10^4$, τ(Lyβ) $\sim 10^3$, τ(Ly8) $\sim 10^2$, and τ(Ly18) ~ 10. For a typical nebula a better approximation than Case A therefore is the opposite assumption, i.e. that it is optically thick in all Lyman lines. This assumption is called *Case B*.

The above two limiting cases have been described in 1938 by Menzel and Baker and are therefore often called Menzel & Baker Case A and Case B. It should not be forgotten that the realistic situation in a nebula is likely in between these two limits. For fairly low lying transitions in the Lyman series Case B will hold, while for the higher transitions, i.e. $n \to \infty$ and small $\tau(\mathrm{Ly}\,n)$, the situation will be more similar to Case A.

18.5 Gas density diagnostics: the emission measure

Let us assume that the fraction of all recombinations that lead through the transition $u \to l$ is p_{ul}. Then, the volume emission coefficient for the line $u \to l$ is (see also Eq. 18.1)

$$\eta_\nu = \frac{h\nu}{4\pi}\,p_{ul}\,\alpha_{\mathrm{A}}\,n_p\,n_{\mathrm{e}}\,\phi_\nu = \eta_\circ\,\phi_\nu \tag{18.22}$$

where we have replaced N^+ with n_p, the number of free protons, for the case of hydrogen. The expression is the multiplication of the number of de-excitations from $u \to l$ per cm^3 per second, per unit solid angle – which is why we need to devide by the total solid angle $\Omega = 4\pi$ – multiplied by the energy of the photon that is emitted $h\nu$, where $\nu = \nu_{lu}$. To get the emitted energy per hz, we need to multiply by the profile function for spontaneous emission ϕ_ν (which has dimension hz^{-1}).

Though to determine the temperature dependent p_{ul} requires a general approach, it is found that $p_{\mathrm{H}\alpha} \sim 0.3$, i.e. only about 30 percent of hydrogen recombinations produce an Hα photon. Similarly, $p_{\mathrm{H}\beta} \sim 0.1$, so only about 10 percent of hydrogen recombinations produce an Hβ photon.

If we observe an H II region along a line-of-sight where there is negligible background specific intensity, and if the emission from the H II region is optically thin, then (see Sect. 4.6)

$$I_\nu^{\rm obs} \simeq \int S_\nu(\tau_\nu)\, d\tau_\nu = \int \eta_\nu(s)\, ds, \tag{18.23}$$

where $I_\nu^{\rm obs}$ is the observed specific intensity and s is the spatial coordinate along the line-of-sight. It follows that the specific intensity integrated over the line proflle is

$$I^{\rm obs} = \int I_\nu\, d\nu = \int \left[\int \eta_\circ\, ds\right] \phi_\nu\, d\nu = \int \eta_\circ\, ds = \frac{h\nu}{4\pi}\, p_{ul}\, \alpha_{\rm A} \int n_p\, n_{\rm e}\, ds, \tag{18.24}$$

assuming the emitting medium has uniform temperature T (recall $\alpha_{\rm A} = \alpha_{\rm A}(T)$). The integral in the last equation is called the *emission measure* of the line-of-sight

$$\mathcal{EM}_H = \int n_p(s)\, n_{\rm e}(s)\, ds. \tag{18.25}$$

If the emitting medium has uniform density, then $\mathcal{EM}_H = n_p\, n_{\rm e}\, D$, where D is the length of the intercept which the line-of-sight makes with the emitting medium.

Thus a measurement of the total line specific intensity Eq. (18.24) enables us to estimate the emission measure of the emitting medium. Suppose, we also have an independent estimate of the linear size D of the emitting region, say from its distance d and its angular size α, via $D = d\,\alpha$. This assumes spherical symmetry, which is often a reasonable approximation. Then we can combine this with the emission measure to obtain $\langle n_{\rm p}\, n_{\rm e}\rangle = \mathcal{EM}_H/D$. Typically in an H II region we have $n_{\rm p} \sim n_{\rm e}$. Therefore

$$\langle n_{\rm e}^2\rangle = \langle n_{\rm p}^2\rangle = \frac{\mathcal{EM}_H}{D}. \tag{18.26}$$

So, from the study of hydrogen recombination lines, we can estimate the density of gas in an H II region and the total mass of ionized gas

$$M_{\rm H\,II} = \frac{4\pi}{3}\left(\frac{D}{2}\right)^3 m_{\rm amu}\, n_{\rm p}. \tag{18.27}$$

18.6 Gas temperature diagnostic: free-free radio continuum emission

How does the continuum spectrum from an H II region look like? As the part of the continuum spectrum that is typically observed is at long wavelengths and the region is almost completely ionized, the extinction is dominated by free-free processes. Therefore, if we look at an H II region with uniform temperature T and the background intensity is negligible, we should observe free-free radiation with specific intensity given by (see Eq. 4.49)

$$I_\nu^{\rm obs} = S_\nu\left[1 - e^{-\tau_\nu}\right] = B_\nu(T)\left[1 - e^{-\tau_\nu}\right]. \tag{18.28}$$

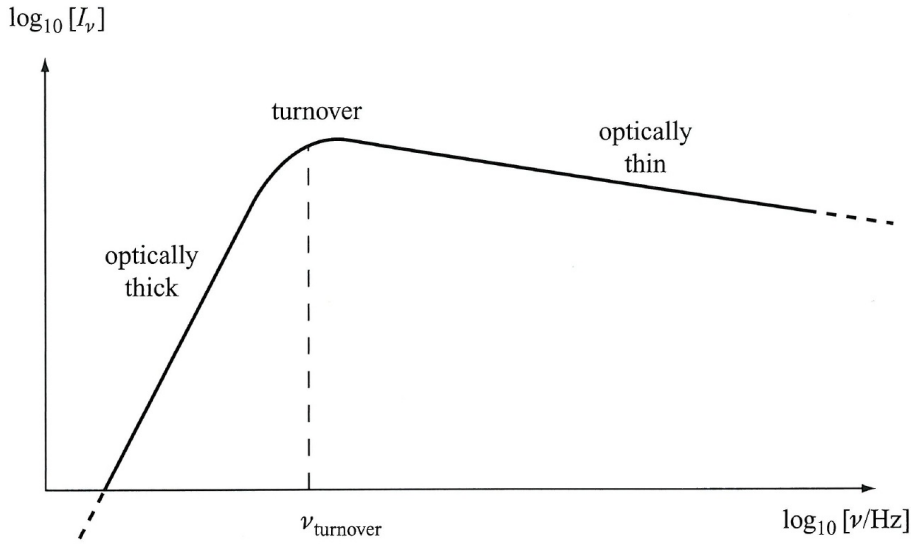


Figure 18.3: *The free-free continuum spectrum from an H II region, showing the transition from optically thick at $\nu \ll \nu_{\rm turnover}$ to optically thin at $\nu \gg \nu_{\rm turnover}$. Figure: Ward-Thompson & A.P. Whitworth (2015).*

Typical H II regions have temperatures $T \sim 10^4$ K and almost always one finds 7 000 K $\lesssim T \lesssim$ 14 000 K. Assuming the medium consists of fully ionized hydrogen only, the free-free optical depth integrated over the H II region is (see Eq. 8.19)

$$\begin{aligned} \tau_\nu &= 1.772 \times 10^{-2}\, \overline{g}_{\rm III}(\nu, T)\, T^{-3/2}\, \nu^{-2} \int n_p\, n_{\rm e}\, ds \\ &= 1.772 \times 10^{-2}\, \overline{g}_{\rm III}(\nu, T)\, T^{-3/2}\, \nu^{-2}\, \mathcal{EM}_H \end{aligned} \tag{18.29}$$

in the Rayleigh-Jeans limit. Recall that in the Rayleigh limit the Planck function reduces to Eq 6.9.

The spectrum of free-free radiation has different asymptotic forms if the medium is optically thick or optically thin. As the optical depth is frequency dependent, we can identify a critical frequency $\nu_{\rm turnover}$ that marks the boundary between these two asymptotic forms of the spectrum. $\nu_{\rm turnover}$ corresponds to the frequency at which the optical depth equals unity.

At high frequencies, the emission is optically thin, so we can write, for $\nu \gg \nu_{\rm turnover}$ and $\tau_\nu \ll 1$

$$I_\nu^{\rm obs} \simeq B_\nu(T)\, \tau_\nu \propto \overline{g}_{\rm III}(\nu, T)\, T^{-1/2}\, \mathcal{EM}_H, \tag{18.30}$$

where the Gaunt factor for free-free processes at long wavelengths is given by Eq. 17.39, scaling approximately as $\overline{g}_{\rm III} \propto \nu^{-0.1}$.

Conversely, at (relatively) low frequencies the emission is optically thick, so we can write, for

$\nu \ll \nu_{\rm turnover}$ and $\tau_\nu \gg 1$

$$I_\nu^{\rm obs} \simeq B_\nu(T) \simeq \frac{2kT\nu^2}{c^2}. \tag{18.31}$$

Figure 18.3 illustrates a typical free-free continuum spectrum from an H II region. One can clearly identify the two asymptotic forms, and the turnover frequency $\nu_{\rm turnover}$.

From the intensity at low frequencies, where the emission is optically thick, we can obtain an estimate of the gas kinetic temperature (see Eq. 6.44)

$$T \simeq \frac{c^2}{2k\nu^2} I_\nu^{\rm obs}, \tag{18.32}$$

for $\nu \ll \nu_{\rm turnover}$.

We can also obtain an estimate of the emission measure from the turnover frequency. The optical depth at the turnover frequency is unity. Hence

$$\mathcal{EM}_H \simeq \frac{56.43}{\overline{g}_{\rm III}(T, \nu_{\rm turnover})} T^{3/2} \nu^2_{\rm turnover} \tag{18.33}$$

Applying Eq. 18.26 we once again have an estimate of the gas density. Remember that these estimates of n_p and $n_{\rm e}$ are strictly speaking root-mean-square values averaged along the line-of-sight.

18.7 Collisional excitation of meta-stable levels and forbidden line emission

The strongest lines in the optical spectrum of many H II regions and planetary nebulae are the *forbidden transitions* in the ground $2p^2$ configuration of O III (see figure 15.1 and 18.4). The transition of the excited ^{1}D$_2$ level to two terms of the ground level, ^{3}P$_1$ and ^{3}P$_2$, produce the eye-catching lines at, respectively, 4959 and 5007 Å. The excitation-potential of the ^{1}D$_2$ level is only 2.51 eV. The highest term of this configuration is ^{1}S at 5.35 eV above the ground level. Other levels of O III have considerably higher excitation potentials. Transitions between levels in the same configuration are called 'forbidden' because they break *Laporte's parity rule* and are denoted by straight brackets, e.g. [O III] λ5007. Forbidden transitions usually have small Einstein coefficients A_{ul}. For [O III]λ4949 en λ5007 for instance these are 0.007 and 0.014 sec^{-1}. This implies that the average lifetime of the ^{1}D$_2$ level of O III is $< t >= 1/0.021 = 36$ sec (zie vgl. 7.31). This is extremely long, therefore these excited levels are referred to as *meta-stable*. Similar conditions as for O III occur for other ions, sometimes with even lower transition probabilities.

Why are the forbidden lines in nebulae so strong compared to allowed transitions? To answer this question we must first consider the formation mechanism of these lines. These lines will certainly not be formed by recombination followed by cascade, as in that case one would have expected both lines from high and low levels. However, save for the O III lines that are caused

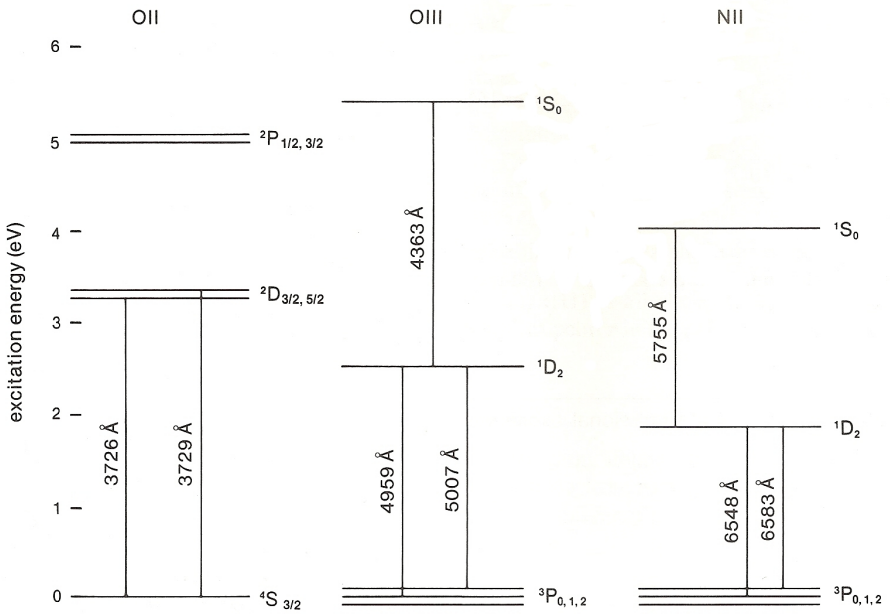


Figure 18.4: *The low-lying energy levels of O* II, *O* III *and N* II. *Excitation energies are relative to the ground level of the ion and are given in electron volt. Figure from Bowers & Deeming (1984).*

by the Bowen fluorescence mechanism (see section 15.4), all O III lines involving higher levels are extremely weak or absent. Note that the meta-stable levels are only a few eV above the ground level, while all other levels that may cause optical transitions need a high excitation energy. The typical energy of a free electron in the plasma is of the order of an electron volt. The Maxwellian velocity distribution (see figure 6.3) shows that in that case there will still be ample numbers of electrons that have energies of a few electron volts, and even some that will have 5 to 7 eV, though energies of say 20 eV will be extremely rare. In other words, there is a reservoir of free electrons that is capable of exciting ions to meta-stable levels. Once excited these ions can fall back through collisional de-excitation or through forbidden line emission.

Now it may be clear why forbidden lines in the optical spectrum of nebulae dominate. It is because the occupation of the meta-stable levels is relatively high. So high even, that it is only a few orders of magnitude below the TE value – given by Boltzmann equation (6.21). The allowed transitions can not use this collisional excitation mechanism as the relevant excitation potentials are too high (typically 10 to 50 eV). These lines can thus only be formed by means of the ionization/recombination/cascade mechanism. However, this process depends on the strongly diluted stellar radiation field (and not exclusively on local conditions as does the collisional process). The number of recombinations (or ionizations) per second is thus many order of magnitude lower than in TE; roughly a factor of W. This difference is so large that even though there are thousands of times as many H II ions than O III ions, the Balmer lines are not capable of dominating the [O III] lines. No, observations show that the green [O III]

lines in nebular spectra can even be ten to twenty times as strong as Hβ.

Gas temperature diagnostics: collisionally excited forbidden lines

The forbidden emission lines of O III at λ4363, 4959 and 5007 Å (see figure 15.1 and 18.4) turn out to be very suited for constraining the temperature of the nebular gas. We assume that the electron density of the gas is much less than the critical density (see section 15.3), such that collisional excitations are negligible.

To keep a clear view of the situation we refer to the ground state $^3\mathrm{P}_{0,1,2}$ as level 1, the first excited level $^1\mathrm{D}_2$ as level 2, and the second excited level $^1\mathrm{S}_0$ as level 3. This notation ignores that there are two[1] possible transitions from level 2 to 1, namely $^3\mathrm{P}_1 - {}^1\mathrm{D}_2$ (λ5007 Å) and $^3\mathrm{P}_2 - {}^1\mathrm{D}_2$ (λ4959 Å), however, we will repair that later. The statistical equilibrium equation (9.19) for level 2 is

$$\begin{aligned} n_2\,A_{21} &= n_1\,C_{12} + n_3\,A_{32} \\ &\simeq n_1\,C_{12}. \end{aligned} \tag{18.34}$$

This reflects that the n_2 level can be populated either by collisional excitations from the ground level ($n_1 C_{12}$) or by spontaneous radiative de-excitations from level 3 ($n_3 A_{32}$). Note that the second process is negligible relative to the first and can be safely ignored, simplifying the derivation. For level 3 it follows that

$$n_3\,(A_{31} + A_{32}) = n_1\,C_{13}. \tag{18.35}$$

This level can be populated by collisional excitations from the ground level ($n_1 C_{13}$) or be de-populated by cascade to level 2 ($n_3 A_{32}$) or level 1 ($n_3 A_{31}$). Note that in this case we do not ignore the process $n_3 A_{32}$ as here it contributes significantly to the depopulation of level 3. The ratio of the populations of levels 2 and 3 is now given by

$$\frac{n_2}{n_3} = \frac{(A_{31}+A_{32})}{A_{21}}\frac{C_{12}}{C_{13}} = \frac{(A_{31}+A_{32})}{A_{21}}\frac{\Upsilon_{12}}{\Upsilon_{13}}\frac{e^{-E_{12}/kT}}{e^{-E_{13}/kT}} = \frac{(A_{31}+A_{32})}{A_{21}}\frac{\Upsilon_{12}}{\Upsilon_{13}}e^{+E_{23}/kT}, \tag{18.36}$$

where we have used the Maxwellian averaged collisional strength $\Upsilon_{lu}(T)$ (see section 7.1) to describe C_{lu}.

In our spectrum we measure the total line flux (eq. 13.6) of each of the three lines. If we assume that each of the lines is optically thin (such that we do not have to solve the equation of transfer and the situation is similar to, for example, that discussed in section 17.4 for the case of Hα emission in stellar winds) the observed line profile integrated flux is given by (see also eq. 7.16)

$$\mathcal{F}_{lu} = \frac{1}{4\pi d^2}\int_V \eta_{lu}\,dV = \frac{1}{4\pi d^2}\int_V \frac{h\nu_{lu}}{4\pi}\,n_u\,A_{ul}\,dV, \tag{18.37}$$

[1] There is a third possibility, namely $^3\mathrm{P}_0 - {}^1\mathrm{D}_2$ (λ4931 Å), that can only occur by means of a quadrupole transition, but that is so weak that it can be ignored altogether.

where d is the distance to the nebula. The line profile integrated emission coefficient η_{lu} (see also Eq. 7.16) multiplied by the total solid angle ($\Omega = 4\pi$) is the total energy emitted in by the spectral line per cm^{-1} per second. It thus needs to be integrated over the entire volume V of the nebular gas. If we assume a constant temperature, the *ratio of the line strengths* 21 and 32 is the same in each cubic centimeter of the nebula and we need not integrate over the full volume to obtain this ratio. We get

$$\begin{aligned}
\frac{\eta_{4959}+\eta_{5007}}{\eta_{4363}} &= \frac{h\nu_{4959}\,A_{4959}+h\nu_{5007}\,A_{5007}}{h\nu_{4363}\,A_{4363}}\,\frac{n_2}{n_3} \\
&= \frac{A_{2321}+A_{4363}}{A_{4959}+A_{5007}}\,\frac{\nu_{4959}\,A_{4959}+\nu_{5007}\,A_{5007}}{\nu_{4363}\,A_{4363}}\,\frac{\Upsilon_{12}}{\Upsilon_{13}}\,e^{+E_{23}/kT} \\
&= \frac{A_{2321}+A_{4363}}{A_{4363}}\,\frac{\overline{\nu}}{\nu_{4363}}\,\frac{\Upsilon_{12}}{\Upsilon_{13}}\,e^{+E_{23}/kT}, \qquad (18.38)
\end{aligned}$$

where $\overline{\nu}$ is the Einstein A averaged frequency of the 21 transition. Notice that we have again split up the 21 transition in the lines $^3\mathrm{P}_1 - {}^1\mathrm{D}_2$ (λ5007 Å) and $^3\mathrm{P}_2 - {}^1\mathrm{D}_2$ (λ4959 Å).

This diagnostic is useful for temperatures ranging from $5\,000 - 20\,000$ K (or $\sim 0.5 - 2$ eV; see eq. 15.7) and earns this sensitivity to the fact that the energy levels 2 and 3 are some distance apart. If indeed the distance between these levels would have been small, $\exp(E_{23}/kT) \sim 1$ and consequently the line flux ratios would have been insensitive to temperature. There are other ions that have a similar favorable positioning of energy levels and that can be used in a similar fashion to constrain the temperature of warm gaseous nebulae, for instance [N II], [O I], [Ne III], [S III]. Figure 18.5 provides an overview of the relevant lines and shows the sensitivity of these diagnostics in the limit of low electron densities.

Equation 18.38 is a good approximation up to $n_\mathrm{e} \sim 10^5\,\mathrm{cm}^{-3}$. At higher electron densities collisional de-excitations start to play a role (see also the discussion on the critical density in section 15.3). The ^{1}D term has a considerably longer lifetime than ^{1}S, and therefore will be de-populated at a lower n_e by collisional de-excitations, causing a weakening of the λ4959 and λ5007 lines. What starts to play a role at electron densities above the value given previously are collisional excitations from ^{1}D to ^{1}S. This strengthens the emission from λ4363. A proper description of this problem requires the solution of the statistical equilibrium equation (9.19), however, an analytical solution that is correct to within first order in $\exp(-\Delta E_{23}/kT)$ is

$$\frac{\eta_{4959}+\eta_{5007}}{\eta_{4363}} = \frac{7.90\exp(3.29\times10^4/T)}{1+4.5\times10^{-4}\,n_\mathrm{e}/T^{1/2}} \qquad (18.39)$$

In the second edition of Osterbrock & Ferland similar approximations are given for [N III], [Ne III], en [S III]. Note that the correction term for the electron density is very small. Even if only a rough estimate of n_e is used, very reasonable estimates for T can be made. Giving eq. (18.39) some thought one must admit that it is a fascinating result: the temperature of the nebular gas can be derived without any knowledge of the local radiation field, the distance to the nebula, and (often) the local electron density.

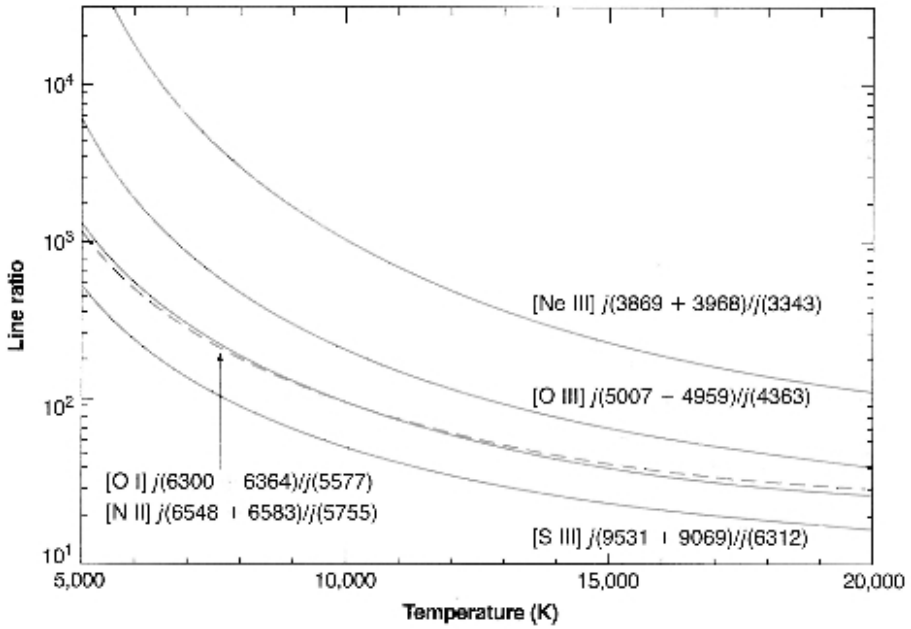


Figure 18.5: *Five temperature sensitive forbidden line ratios as function of the electron temperature. The [O I] (solid line) and [N II] (dashed line) are almost superimposed, partly because they have very similar excitation potentials. All ratios shown are in the limit* $n_{\rm e} \ll n_{\rm e}^{\rm crit}$ *(*$n_{\rm e} = 1\,cm^{-3}$*). Figure from Osterbrock & Ferland (2006).*

Gas density diagnostics: fine-structure emission lines

An estimate of the electron density $n_{\rm e}$ in the rarefied nebular gas can be made by determining the ratio of the line strengths of forbidden transitions in the ground configuration of ions with very comparable excitation energy (contrary to what is required for a good T-diagnostic, see the above discussion on [O III]). The two best examples of such a situation are [O II] $\lambda3729/\lambda3726$ en [S II] $\lambda6716/\lambda6731$[2].

We again take oxygen as an example. To keep a clear view of the situation we refer to the ground level $^4{\rm S}^\circ_{3/2}$ as level 1, the first excited level $^2{\rm D}^\circ_{5/2}$ as level 2, and the second excited level $^2{\rm D}^\circ_{3/2}$ as level 3. The two highest terms in the ground level configuration, $^2{\rm P}^\circ_{1/2,3/2}$, need not be considered. In formulating the relevant statistical equilibrium equations (9.19) we consider collisional excitations, collisional de-excitations and spontaneous de-excitations. Note that we ignore the (forbidden) transition $^2{\rm D}^\circ_{3/2} - {}^2{\rm D}^\circ_{5/2}$. Though this is a good approximation for the radiative transition between these two fine structure levels it is not so for the collisional coupling. However, in the two limiting situations $n_{\rm e} \to 0$ and $n_{\rm e} \to \infty$ this is not a problem: in the first instance these collisions are indeed negligible; in the second instance the

[2]Notice that for the oxygen line the line ratio is longest wavelength / shortest wavelength of the doublet while this is reversed in case of sulpher. This is because in these ions the $^2D^\circ_{3/2,5/2}$ are interchanged, see figure 18.7.

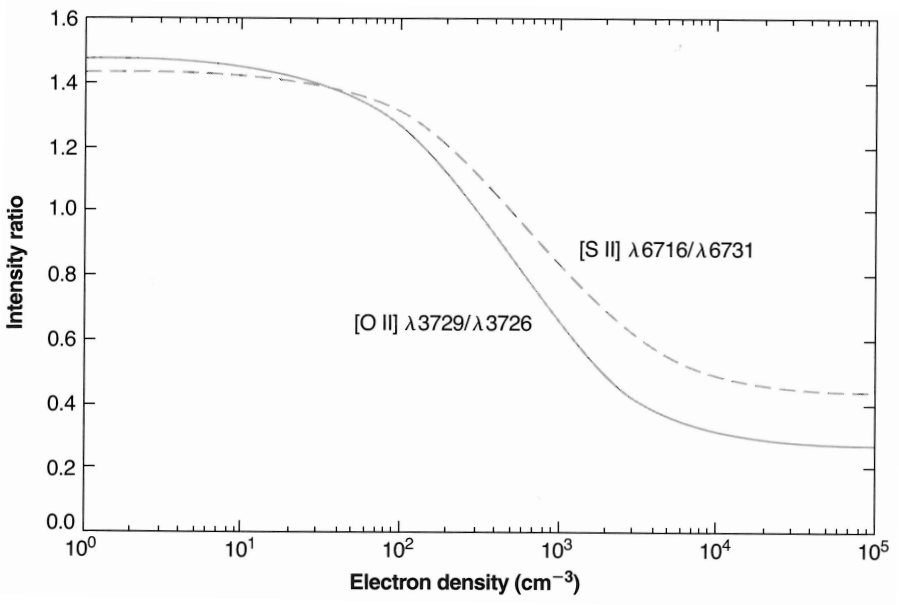


Figure 18.6: *Statistical equilibrium calculation of line ratios for [O II] (solid line) and [S II] (dashed line) as a function of $n_{\rm e}$ for a temperature $T = 10\,000$ K. At other temperatures the relations shown are nearly correct if one assumes the horizontal scale represents $n_{\rm e}(10^4/T)^{1/2}$.*

two levels (2 and 3) will be in LTE relative to each other. This is automatically taken care of by considering the collisional transitions between $1-2$ and $1-3$. The situation of arbitrary $n_{\rm e}$, however, do require a consideration of the ${}^2{\rm D}^\circ_{3/2} - {}^2{\rm D}^\circ_{5/2}$ coupling.

We get for level 2 is

$$n_2\,(A_{21} + C_{21}) = n_1\,C_{12}, \tag{18.40}$$

and for level 3

$$n_3\,(A_{31} + C_{31}) = n_1\,C_{13}. \tag{18.41}$$

For the ratio between the line strengths 21 and 31 it follows that

$$\begin{aligned}\frac{\eta_{21}}{\eta_{31}} = \frac{\eta_{3729}}{\eta_{3726}} &= \frac{n_2 A_{21} h\nu_{21}}{n_3 A_{31} h\nu_{31}} \simeq \frac{A_{21}}{A_{31}}\frac{C_{12}}{C_{13}}\frac{(A_{31}+C_{31})}{(A_{21}+C_{21})} = \frac{A_{21}}{A_{31}}\frac{C_{12}}{C_{13}}\frac{C_{31}}{C_{21}}\frac{(A_{31}/C_{31}+1)}{(A_{21}/C_{21}+1)} \\ &\simeq \frac{A_{21}}{A_{31}}\frac{g_2}{g_3}\frac{(A_{31}/C_{31}+1)}{(A_{21}/C_{21}+1)}.\end{aligned} \tag{18.42}$$

The approximately equal sign denotes that $\nu_{21} \simeq \nu_{31}$. For the last equality we have used eq. (7.9) and realized that $E_{12} \simeq E_{13}$.

In the low density limit ($n_{\rm e} \to 0$), each collisional excitation is followed by the emission of a photon. One obtains $\eta_{3729}/\eta_{3726} = C_{12}/C_{13} = \Upsilon_{12}/\Upsilon_{13}$. In the high density limit, that for ${}^2{\rm D}^\circ_{5/2}$ is reached at $n_{\rm e}^{\rm crit} \sim 3 \times 10^3\,{\rm cm}^{-3}$ and for ${}^2{\rm D}_{3/2}$ at $n_{\rm e}^{\rm crit} \sim 1.6 \times 10^4\,{\rm cm}^{-3}$, it follows that $\eta_{3729}/\eta_{3726} = A_{21}\,g_2/A_{31}\,g_3 = 0.34$.

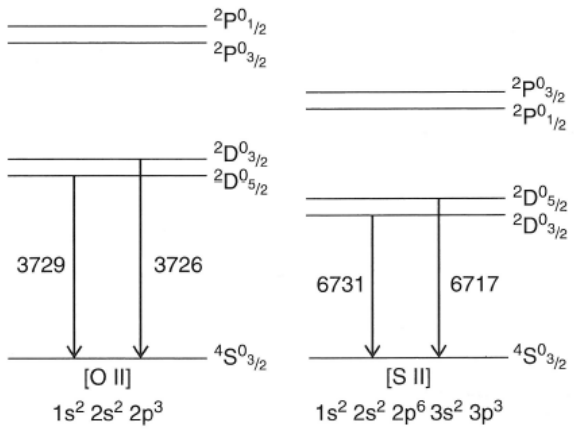


Figure 18.7: *Energy-level diagram of the $2p^3$ ground configuration of O II and $3p^3$ ground configuration of S II, relevant for the formation of forbidden [O II] and [S II] lines. Note that the fine-structure levels $^2D^\circ_{3/2,5/2}$ and $^2P^\circ_{1/2,3/2}$ (the energy separation of the different J-levels is exaggerated for clarity) are switched around in these two ions.*

Figure 18.6 shows the behavior of η_{3729}/η_{3726} as function of n_e for the exact solution of the statistical equilibrium equations, also accounting for collisional excitations to the $^2\mathrm{P}^\circ_{1/2,3/2}$ levels.

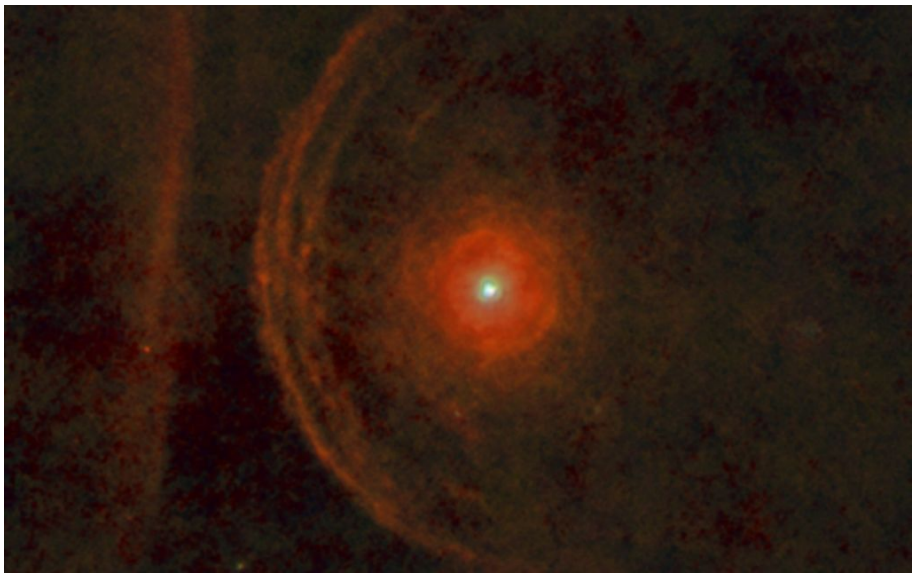


Figure 18.8: *Composite color image of the Herschel PACS 70, 100, 160 μm images of the Red Supergiant Betelgeuse (α Orionis) and its surroundings. Betelgeuse travels through its local intersteller environment at a speed of 25 $\mathrm{km\,sec^{-1}}$, plowing up material forming an asymmetric arc or bow shock. The apex of the arc is at 5.75' from the star. Notice to the left a dust filament, thought to be unassociated with the bow shock. From: ESA-Herschel-PACS / Decin et al. 2012.*

Exercise 18.1

The Red Supergiant star Betelgeuse in the constellation of Orion travels through its local interstellar environment at a speed of $v_\star$ = 25 $\mathrm{km\,sec^{-1}}$, creating a bow shock. The apex of this arc-like structure has a radius of 5.75 arcmin from the star which is at 200 pc. We assume the bow shock is observed edge on.

a) Compare the apparent apex radius in arcmin to the apparent radius (also in arcmin) of the moon.

b) Compute the apex radius in pc.

For a bow shock, the *standoff distance* R_{so} is where the ram pressure of the stellar wind of the moving star balances that of the surrounding material. R_{so} is measured in the direction of motion of the star. The mass-loss rate is given by

$$\dot{M} = 4\pi\, r^2\, \rho_{\mathrm{w}}(r)\, v_{\mathrm{w}}, \tag{18.43}$$

where r is radial distance, ρ_{w} the wind density and v_{w} the wind velocity. For Betelgeuse $v_{\mathrm{w}} = 17\,\mathrm{km\,sec^{-1}}$. We assume the ambient medium to be a neutral gas, for which the mean molecular weight is $\mu = 1.27$, and adopt a canonical value for the particle density $n_\circ = 1\,\mathrm{cm^{-3}}$.

c) Derive a formula for the mass-loss rate of Betelgeuse as a function of $R_{\rm so}$ and other relevant quantities.

d) Calculate the mass-loss of Betelgeuse in solar masses per year. The literature value is $\dot{M} = (3 \pm 1) \times 10^{-6} M_{\odot} \mathrm{yr}^{-1}$. How well does your result compare to this value? Which assumption(s) may contribute to the cause of the difference.

19

ISM and IGM

The medium in-between the stars of our Milky Way is filled with interstellar matter. About 99 percent of the mass of this material is due to gas and about 1 percent due to solid state particles. The gas may be in an atomic, ionized, or (ionized) molecular state. The typical sizes of interstellar solid state particles are 0.01 to 0.1 μm, although grains up to a few micron in diameter may be present as well. Though only responsible for a very small fraction of the mass of interstellar matter, it is the extinction properties of these dust particles that most affect the stellar light traveling through interstellar space. We briefly discuss the impact of interstellar gas on the appearance of the spectra of stars. Most of the chapter, however, concentrates on the effect that interstellar dust can have on the stellar energy distribution. We will not care much about the origin and nature of this dust, nor the processes that shape its properties.

19.1 Introduction

Interstellar and intergalactic absorption lines

Stellar spectra may contain lines that originate from material in the interstellar medium. These *interstellar absorption lines* become stronger in more distant stars as the result of superpositioning of an increasing number of interstellar clouds. When the spectral resolution and signal-to-noise are high enough to resolve these multiple structures one often finds differences among lines of different elements, demonstrating that there are differences between clouds.

The interstellar lines originate from ground-level (resonance) transitions. Examples of well known interstellar lines are Lyα λ1216, C IV $\lambda\lambda$1548,1551, Mg II $\lambda\lambda$2796,2803 in the ultraviolet part of the spectrum, and Ca I λ4227, Mg I λ4571, and Na I $\lambda\lambda$5890,5896 in the optical. Interstellar lines may also be due to molecules, such as, in the optical, NH, CH, CH^+, CN and C_2.

It is not always trivial to distinguish interstellar lines from the, often, plentiful stellar lines. Hints that point to lines originating in the interstellar medium, rather than in the stellar atmo-

sphere or in a strictly circumstellar region are:

- Certain absorption lines are seen in the spectra of distant stars, but are not seen in the spectra of otherwise very similar nearby stars.
- Certain absorption lines seen in stellar spectra are due to species which are not anticipated to be present in the atmosphere of the background star. Usually they are species expected to be in a lower or higher stage of ionization or to be dissociated at the temperatures and densities of the background source.
- Certain absorption lines seen in stellar spectra are too narrow to have been produced in a stellar atmosphere. The thermal width of a spectral line is given by Eq. 13.19. The typical temperatures in stellar atmospheres are much higher than those of the bulk of the interstellar medium, that is rather cold. Additional broadening may cause the stellar lines to be even wider. Perhaps the easiest way to separate the interstellar lines from the stellar lines is to observe stars that are rotating rapidly enough to make their lines wash out (see Sect. 13.3).
- The central frequency of an absorption line may indicate that the absorbing gas has a different radial velocity from the background star. The existence of interstellar gas was first unambiguously confirmed from observations of a spectroscopic binary. The central frequencies of the absorption lines produced in the stellar atmosphere shifted with a regular period, as the star orbited its companion. The central frequencies of the absorption lines produced in the intervening interstellar medium were constant.
- Often close groups of absorption lines are observed, with all the lines in a group being attributable to the same transition in the same species. The inference is that the absorbing particles are not distributed uniformly along the line of sight to the background star, but are concentrated in discrete clouds having different bulk radial velocities. The number of lines in a group tends to be larger for more distant stars as, on average, the number of intervening clouds increases with the distance to the background star.

The most prominent intergalactic absorption line is Lyα $\lambda 1216$. Fig. 19.1 shows a high-resolution spectrum ($R \sim 37\,500$) of the quasar QSO Q1422+231. This active galactic nucleus produces strong emission in Lyα centered at 5622 Å, from which a redshift $z = 3.625$ is derived. At shorter wavelengths the emission is eaten away by a great number of sharp absorption lines. Most of these are single Lyα lines formed in galactic halos and gas clouds located between the quasar and us. They appear at many discrete redshifts between $z = 3.625$ and 0 and are collectively known as the *Lyα forest*. They are an important probe of physical conditions in galaxies and the ISM at early times in the evolution of the universe.

Discrete Interstellar Bands

In the optical part of the spectrum there are over a hundred rather weak absorption bands. Though these bands are already known since the 1930s, their carriers are still not identified.

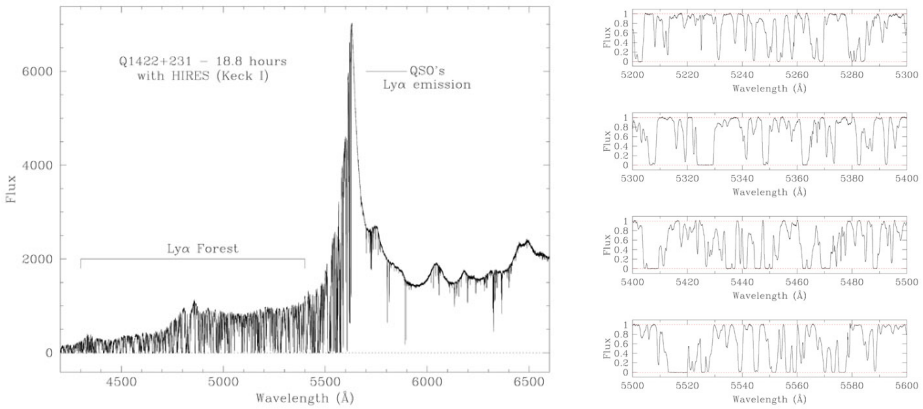


Figure 19.1: *Left: Spectrum of the high redshift quasar QSO Q1422+231, obtained with the High Resolution Echelle Spectrograph on the Keck telescope in Hawaii. The strong and broad Lyman α emission at 5622 Å is produced in the galaxy itself and indicates that the quasar is at a redshift $z = 3.625$. The numerous Lyman absorption lines at larger wavelengths are produced by intervening galaxies and clouds and are collectively referred to as the Lyman α forest. Right: detail of the Lyman α forest. Each 100 Å wide stretch of continuum normalized spectrum contains more than 50 individual absorption components. From: Ellison (2000).*

The bands are known as the *Diffuse Interstellar Bands* or DIBs. The wavelengths of the DIBs range from the blue (for instance the strongest one at 4428 Å) to the far red (e.g. the one at 8650 Å). There are indications for different sets of bands. This may imply that not all bands are caused by the same material. In many cases the strength of the diffuse bands are correlated with the amount of interstellar absorption by dust. This suggests that the material that is responsible for the interstellar absorption is also responsible for the diffuse bands, or that the material that caused the DIBs is well mixed with the interstellar dust. Surprisingly enough, there is no evidence for the presence of DIBs in *circumstellar dust*, i.e. dust in the direct surroundings of a star, or in comets. Apparently the diffuse bands are a purely interstellar phenomenon.

The prevailing suspicion at the moment is that the carriers are complex, carbon-based (i.e. 'organic') molecular structures, possibly Poly-Aromatic Hydrocarbons (PAHs); they are probably ionized and perhaps protonated. Strangely enough, PAHs features *are* seen in spectra of circumstellar envelopes and cometary tails, where DIB features are conspicuously absent.

Interstellar dust

The existence of solid state material between the stars was first proposed by Otto Wilhelm von Struve in 1847 based on the analysis of star counts which suggested that the number of stars per unit volume decreased with increasing distance from the sun. Struve proposed that the starlight was experiencing absorption proportional to distance. It was not until 1909 that

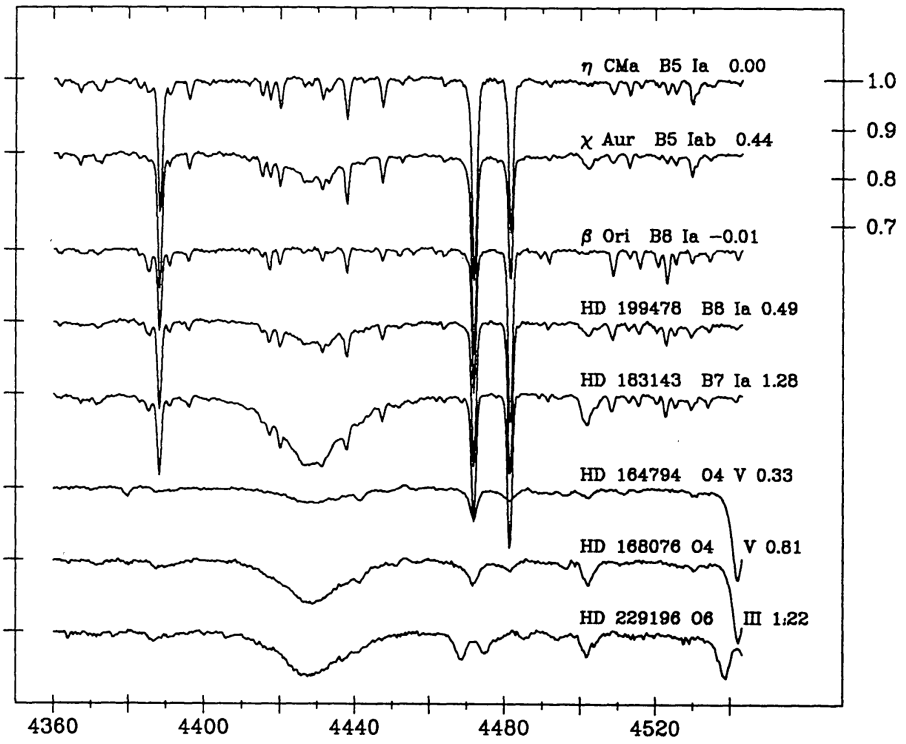


Figure 19.2: *The spectral region containing the diffuse band at 4428 Å in both weakly and strongly reddened stars. Spectral types and $E(B-V)$ values appear on the right. All spectra have been shifted in wavelength so that their stellar lines coincide. For similar spectral type and luminosity class unreddened and reddened spectra have been included. The broad λ4428 feature is more apparent in the reddened stars, as is the DIB at λ4501, although the latter is blended with stellar Ti II lines in the B stars. The lowermost star, HD 229196, is a double-line spectroscopic binary with a line splitting of about 400 $\mathrm{km\,sec^{-1}}$. From: Herbig, G.H., 1995, ARA&A 33, 19.*

Jacobus Cornelius Kapteyn realized the full significance of this interstellar extinction. Shortly thereafter Barnard documented the irregular variations in the distribution of the absorbing matter. The identification of small solid state particles as the source of this extinction was finally accepted in the 1930s through the work of Trumpler and Stebbins, Huffer, and Whitford.

The properties of the interstellar particles can be studied through their emission, scattering, and absorption properties. The most abundant dust particles are composed of either amorphous silicates or carbon. The latter probably is in the form of graphite. The silicates have a chemical composition that is about that of olivine ($\mathrm{Mg}_x\mathrm{Fe}_{2-x}\mathrm{SiO}_4$ where $0 \leq x \leq 1$) and can be spectroscopically identified by emission or absorption bands at ~ 9.7 and 18 μm, though these grains absorb quite effectively over a wavelength interval that extends from the ultraviolet to the far-infrared. The graphite grains too show this broad continuum absorption and are likely responsible for an absorption band at 2175 Å. In regions where the interstellar particle density is relatively high, i.e. in molecular clouds, water ice may freeze on the dust particles,

causing an absorption at $\sim 3.08\ \mu\mathrm{m}$. Absorption by solid carbon monoxide at $\sim 4.08\ \mu\mathrm{m}$ is often observed in molecular clouds.

19.2 Continuum extinction by interstellar dust

We consider the line of sight towards a star that emits an intensity $I_\lambda^\star$. If the optical depth in the ISM in between us and the star is τ_λ, the observed intensity will be given by

$$I_\lambda(0) = I_\lambda^\star e^{-\tau_\lambda} \tag{19.1}$$

The amount of extinction of starlight is often expressed in magnitudes. If we introduce

$$A_\lambda \equiv m - m_\circ \tag{19.2}$$

where $m_\circ$ is the intrinsic magnitude of the stellar light, we find for the relation between extinction and optical depth

$$A_\lambda = -2.5 \log \left[\frac{I_\lambda(0)}{I_\lambda^\star} \right] = 1.086\, \tau_\lambda \tag{19.3}$$

This equation shows that the extinction of starlight by dust expressed in magnitudes is almost equal to the optical depth (at the chosen wavelength).

Interstellar extinction law

The wavelength dependence of the interstellar extinction in the direction of three stars is shown in figure 19.4, where A_λ/A_V is the interstellar extinction normalized to the photometric V band. The functional behaviour of this quantity is refered to as the *interstellar extinction law*. It only depends on the (mean) intrinsic properties of dust particles in the beam towards the star, and not on the length of the beam. The use of A_V, the extinction in the V band (see § 6.6) to normalize the extinction law is arbitrary and one could argue that it is more meaningful to use, for instance, the extinction in the K filter (centered around $\sim 2.14\ \mu\mathrm{m}$) as for these wavelengths the extinction is almost independent of the direction in which we look. Detailed analysis of the extinction law in many directions shows that the behaviour of A_λ/A_V can well be described with only one free parameter, the *total-to-selective extinction*

$$R_V \equiv \frac{A_V}{A_B - A_V} = \frac{A_V}{E(B{-}V)} \tag{19.4}$$

The *selective extinction* is formally defined as $E(\lambda_1 - \lambda_2) \equiv A_{\lambda_1} - A_{\lambda_2}$ (see also eq. 19.6). Observed values of R_V are inbetween ~ 2.75 and 5.5. As this number is always positive, it implies that the extinction in the B band is always larger than that in the V band. This can also be seen in figure 19.4. We will return to this below. For dust particles in the diffuse

Filter	$\lambda\,[\mu\mathrm{m}]$	A_λ/A_V
U	0.36	1.569
B	0.44	1.337
V	0.55	1.000
R	0.70	0.751
I	0.90	0.479
J	1.25	0.282
H	1.65	0.190
K	2.20	0.114
L	3.45	0.056

Figure 19.3: *A_λ/A_V at optical and near-IR wavelengths for $R_V = 3.1$ following Cardelli et al. (1989).*

(read: typical) interstellar medium one typically derives $R_V \sim 3.1$. Larger values of R_V are found in lines of sight towards dense molecular clouds. A larger value for R_V implies a decrease of A_B/A_V. This is equivalent to an increase of the "greyness" of the extinction. A grey(er) extinction is characteristic for relatively large dust grains, and the tendency to find relatively large values of R_V in directions in which the interstellar medium has a larger density therefore strongly suggests that the dust particles in these clouds are relatively large. How large is relatively large? The behaviour of the extinction of light on solid state particles is dependent on the value $2\pi a/\lambda$, where a is the radius of the grain. For $2\pi a/\lambda \gg 1$ the extinction is grey as the cross section of the particle reaches its geometrical surface πa^2; for $2\pi a/\lambda \lesssim 1$ the extinction properties are complex, however usually they are proportional to λ^{-2} to λ^{-1}; only for $2\pi a/\lambda \ll 1$ the wavelength dependence becomes that of Rayleigh scattering, i.e. λ^{-4}. Relatively large therefore implies $a \sim 1$–$2\ \mu\mathrm{m}$. The "growth" of dust particles occurs through coagulation, i.e. the sticking of grains in collisions.

The strongest spectroscopic feature in the extinction curve is a conspicuous "bump" at 2175 Å or 4.6 $\mu\mathrm{m}^{-1}$. This feature is visible for all values of R_V. Its origin is not well understood. In view of the prominent nature of the bump it must be formed by an abundant material. Graphite, an ordered and stable form of carbon, has a strong resonance at 2175 Å with about the correct width and strength to explain the bump. While identification of the 2175 Å bump is still not certain, many ascribe it to graphite or perhaps a somewhat less ordered form of carbon.

Cardelli, Clayton, and Mathis (1989) in a seminal paper derived an empirical expression for the behavior of A_λ/A_V. These ratios are given in table 19.3 for wavelengths centered on photometric passbands. In the near-infrared wavelength range between 0.9 and $3.3\,\mu\mathrm{m}$ − where the photometric filters I (0.9 μm), J (1.25 μm), H (1.65 μm), and K (2.20 μm) are positioned − the extinction can be well described by a power law

$$\frac{A_\lambda}{A_K} = \left(\frac{\lambda}{2.20\,\mu\mathrm{m}}\right)^{-1.61} . \tag{19.5}$$

This behavior implies, for instance, that $E(J-K) = 1.52\,E(H-K)$.

Unlike the large differences from sightline to sightline in the ultraviolet and optical part of the spectrum, the relative extinction properties in the near-IR are the same irrespective of the chosen line of sight. The universal nature of the extinction law in the near-IR implies that the size distribution of the largest grains is similar in all directions. For wavelengths larger than $\sim 5\ \mu\mathrm{m}$ the extinction is not well known. Beyond a wavelength of 20 μm the uncertainty in the extinction is probably a factor of two. Realize that this far into the infrared the extinction is often negligible anyway.

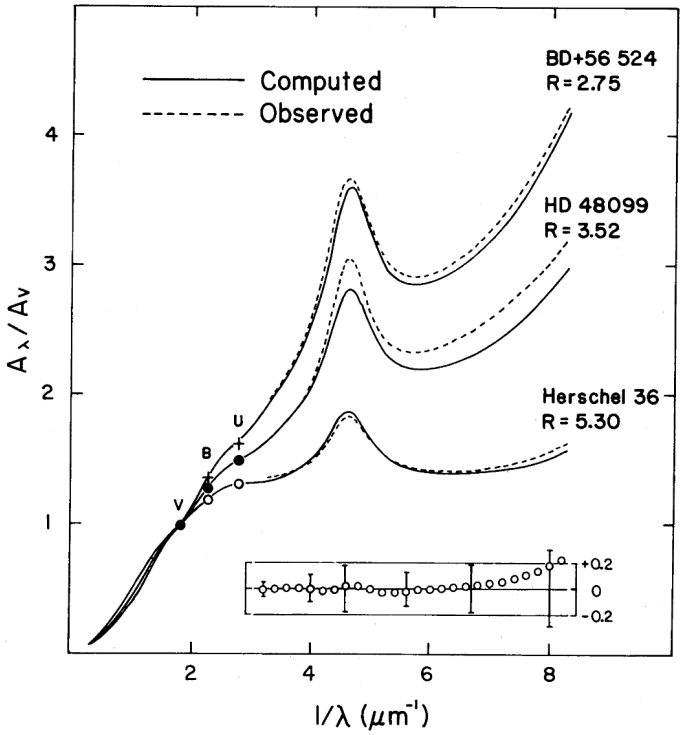


Figure 19.4: *The measured wavelength dependence of the interstellar extinction for three directions into the ISM (dashed lines). The extinction is normalized to the value in the V band (5500 Å). The extinction in the direction of Herschel 36 – the ionizing star in the H II region M8 – is considered "peculiar". The full lines are fits to the measured A_λ/A_V. From: Cardelli, Clayton & Mathis (1989).*

Interstellar reddening

We saw that at wavelengths larger than 2175 Å the extinction decreases with increasing wavelength. Blue starlight therefore suffers more from dust in the ISM than does red starlight, causing a *reddening* of the spectral energy distribution emitted by the star. The amount of reddening is often expressed in terms of the visual selective extinction or *color excess*

$$E(B{-}V) = A_B - A_V = (B-V)_{\rm obs} - (B{-}V)_\circ \qquad (19.6)$$

where $(B{-}V)_{\rm obs}$ is the measured and $(B{-}V)_\circ$ the intrinsic color of the star. In our Milky Way the interstellar dust is concentrated in the galactic plane, with an effective scaleheight of about 100 pc. The mean reddening in the plane is about 0.61 magn kpc^{-1}. If we take $R_V = 3.1$ we get an $A_V = 1.9$ magn kpc^{-1}. Beware that in reality the distribution of dust is very patchy, i.e. concentrated in small and large interstellar clouds, and that there are directions in which the reddening deviates a factor of 5 to 10 from the mean. Measurements of the correlation between the column density of gas and the interstellar reddening yield an average gas-to-color-excess ratio N(H I + H_2) / $E(B{-}V)$= 5.8 10^{21} atoms cm^{-2} magn^{-1}, where N(H I +

H_2) = N(H I) + 2N(H_2). For atomic hydrogen only the mean value is N(H I) / $E(B-V)$= 4.8 10^{21} atoms cm^{-2} magn^{-1}.

In view of the distribution of dust in our Milky Way interstellar reddening is particularly problematic for stars located in the galactic plane, i.e. population I stars, and that are at reasonably large distances. However, also (young) stars in star forming regions can suffer from a large extinction due to dust grains in the molecular clouds from which these stars have recently (up to a few million years ago) formed. Part of the extinction of starlight in star forming regions can be caused by *circumstellar dust*, i.e. dust in the direct surroundings of a young star. The extincion curve of this circumstellar dust can deviate strongly from that of ISM dust. The reason for this is that the proximity of the central star and the potential high density of the circumstellar material can lead to chemical alteration, crystallization or coagulation of the grains. This affects the extinction properties. A different category of stars that may suffer strongy from circumstellar extinction are red giants and supergiants. In the stellar winds of these cool stars dust may actually condense. The (super)giants with extremely strong stellar winds ($\dot{M} \sim 10^{-4}$ à 10^{-3} $M_\odot$yr^{-1}) can form such large amounts of circumstellar dust that they can no longer be observed in the optical – though they can still be observed in the infrared. One may expect that the extinction curves of these so-called OH/IR stars differ strongly from that of ISM dust.

Reddening in the color-color diagram

For the typical interstellar medium, which has R_V = 3.1 magn, the color excess $E(U-B)$ scales linearly with $E(B-V)$, following

$$E(U-B) = 0.72\,E(B-V) \tag{19.7}$$

Consequently the quantity

$$Q = (U-B) - 0.72(B-V) \tag{19.8}$$

for any star is independent of the reddening in the direction of the star, provided the star suffers no additional circumstellar extinction.

Affection of the spectrum by interstellar reddening

The effect of interstellar reddening on the color indices $U-B$ and $B-V$ is shown in figure 19.5. As an example we show the effect that corresponds to a visual extinction $A_V = 1$. The open symbols denote the intrinsic colors; filled symbols are for reddened colors. If, say, the only thing we have available for a star are photometric measurements in UBV, and if we would assume that interstellar extinction is negligible, we would, for instance, wrongly classify an F2 V star as a G5 V star if in reality $A_V = 1$ magn. Phrased differently, if one erroneously neglects interstellar reddening one will underestimate the effective temperature of the star.

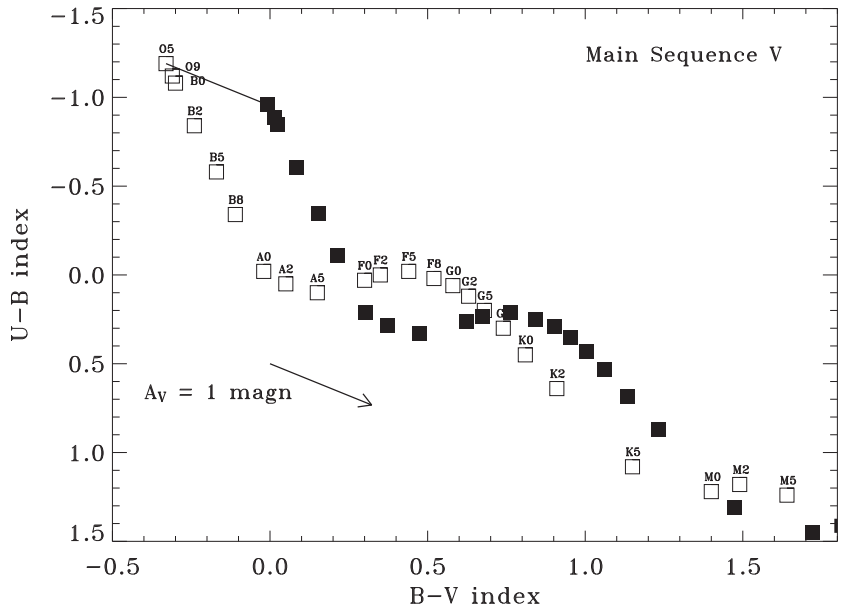


Figure 19.5: *Relation between the color indices $U-B$ and $B-V$ for dwarf stars (luminosity class V). The open symbols denote the intrinsic colors; the filled symbols are the observed colors for an interstellar extinction $A_V = 1$ magn. To a certain extent, one could say that interstellar reddening simulates a lower $T_{\rm eff}$.*

Figure 19.6 shows the effect of interstellar reddening on the spectrum of a G5 V star. In the top panel the flux is plotted on a linear scale, such that it is clearly visible that for A_V = 1 magn the observed optical flux at 5500 Å (i.e. in the V band) is only $\exp(-1/1.086) \sim 0.4$ times the intrinsic value, i.e. the starlight at this wavelength is weakened by a factor 2.5. The lower panel (in which the flux is given on a logarithmic scale) again shows that due to a more efficient extinction at short wavelengths the stellar spectrum will be reddened, and that if the extinction is very large only stellar light at infrared or even longer wavelengths can be observed.

There are different ways to measure the amount of interstellar extinction towards a star. In all cases it is important that the optical spectrum of the star is available, such that its spectral type, and, ideally, its luminosity class can be determined.

- If the spectral type and luminosity class are known, the intrinsic $(B-V)_\circ$ can be obtained from table B.5. Measurement of $B-V$ then yields $E(B-V)$ (see eq. 19.6). This method does not constrain the value of R_V.

- A Kurucz or other model, or an observed spectrum of a nearby non-reddened star that is characteristic of the spectral energy distribution of the star of interest (to be determined

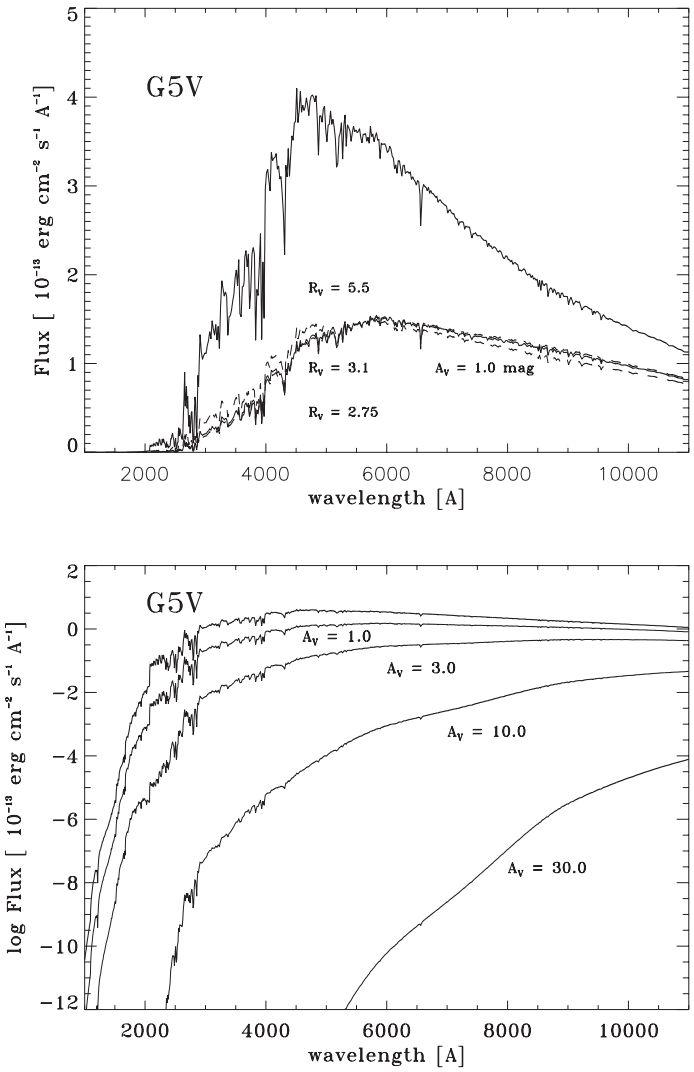


Figure 19.6: *Examples of the effect of interstellar reddening on the energy distribution of a G5 V star. The top panel shows the flux on a linear scale; the bottom panel on a logarithmic scale. To describe the extinction curve we adopted the fit formula by Cardelli et al. (1989) using* $R_{\rm V}$ *= 3.1.*

from the known spectral type & luminosity class), is artifically reddened and fitted to the measured UBVRI broadband photometry. The interstellar extinction can e.g. be described by the fit formula given by Cardelli et al. (1989). In principle, this method can also provide some information on the value of R_V.

- A measured spectrum, from the UV to the near-IR, is divided by a representative intrinsic spectrum (observed or predicted – see previous method). This provides detailed

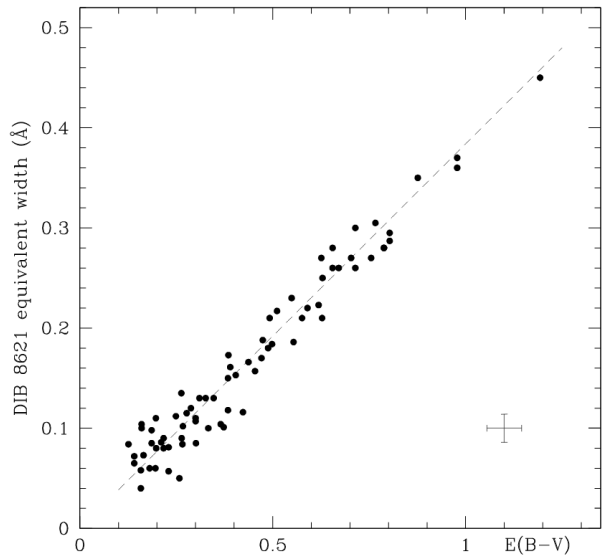


Figure 19.7: *The equivalent width of the diffuse interstellar band at 8620.4 Å as function of reddening. The dashed lines is the best fit,* $E(B-V) = 2.72(\pm 0.03) \times W_{8620}$. *The typical error bar on the individual points is given in the lower right corner. From: Munari et al. (2008).*

information on the interstellar extinction curve in the line of sight towards the target star (see eq. 19.3).

- An independent measurement of extinction may be obtained from the strength of the diffuse interstellar bands. As an example, figure 19.7 shows the correlation between the strength of the DIB at λ8620.4 Å and reddening in our Galaxy.

19.3 Dust in and in-between galaxies

Dust not only resides in the interstellar medium of our Milky Way, but is a property of galaxies in general. Just as the dust in our galaxy betrays its presence by causing dark silhouette structures (by blocking the light from stars located behind these structures), dust in other galaxies also produces such silhouettes (see figure 19.8). If we observe individual stars in other galaxies, and we correct for dust in our Milky Way, we also see the wavelength dependent extinction that is so characteristic for small solid state particles. A third way to prove the presence of interstellar dust in other galaxies is to focus on the thermal emission that this material produces. Heated by the stars in the galaxy, or by non-thermal radiation from the central parts of the system, dust will emit at infrared wavelengths. This last detection method of dust in other galaxies has shown itself to be especially succesful in the study of distant galaxies, that can not be resolved spatially.

Figure 19.8: *The irregular galaxy NGC 5128 (Centaurus A), at a distance of 4.6 Mpc, shows a dark band that is the result of the absorption of starlight by interstellar dust grains present in the system. The dust lane is possibly the remnant of the merging of this system with a smaller spiral galaxy, a few billion years ago. The collision also caused an explosive increase in the star forming activity (a so-called* starburst*). From: ESO*

The study of the dust extinction laws in the Magellanic Clouds is especially intriguing as the interstellar gas in both these satellite galaxies has elemental abundances that are different from those in our Galaxy. For instance, the abundance of oxygen in the Large Magellanic Cloud is about a factor of two lower than in our Milky Way, while that in the Small Magellanic Cloud is even lower by about a factor of five. Such differences in the chemical abundance pattern of the interstellar gas may hint to possible differences in the interstellar dust population, and hence in the extinction laws of our two satellites. It is found that the shape of the extinction law is strongly dependent on the line of sight that is probed (even within the LMC or SMC). For the Large Magellanic Cloud some lines of sight show a similar extinction curve as derived for our galaxy. But there are also directions within the LMC for which the bump in the extinction around 2175 Å is weaker and the rise in extinction to the far ultraviolet is steeper. In the case of the SMC the extinction law lacks the 2175 Å bump. For this system the extinction law in the near infrared, optical, and ultraviolet is reasonably well represented by a single power law.

Is there *intergalactic dust*, i.e. are there solid state particles residing in the space between

galaxies? The existence of such dust could have severe consequences for, for instance, the cosmological distance scale, and potentially be an important source of dark matter. So far, however, it has not been found. The search for any intergalactic dust can – as in the case of dust in the ISM – be conducted in three ways: by looking at effects of extinction, obscuration, or infrared emission. There might be dust grains between galaxies that are part of a cluster of galaxies. It would be relatively easy to look for this *intracluster dust* (that following our definition should also be classified as intergalactic dust) if only for the fact that one knows where to look for it. One possibility is to investigate whether distant galaxies seen through a nearby cluster seem fainter and/or redder compared to distant galaxies in lines of sight that do not cross the cluster. Alternatively, one could use the idea that extinction by intracluster gas may render distant background galaxies so faint that they are no longer visible. If so, one would detect a deficiency of faint galaxies in the direction of nearby clusters.

If intergalactic dust truly exist, it may lead to a relation between intrinsic brightness and/or color of galaxies as a function of distance. For example, do quasars of higher redshift appear redder than those of lower redshift? A complicating problem in detecting dust grains in intergalactic space is disentangling the effects of such IGM dust from evolutionairy effects in the systems used as probes. After all, changes in the colors of galaxies on very long timescales (billions of years) may also be caused by changes in, for instance, the stellar population and/or amount of interstellar dust in these systems. An alternative method to detect IGM dust would be to use the scattering properties of dust grains at röntgen (*X-ray*) wavelengths. Quasars, for instance, are very bright in X-rays and should possess detectable X-ray halo's if the intergalactic medium contains a sizeable amount of grains. Such haloes have yet to be searched for. One might also study the distribution of galaxies on the sky: are there "holes" in this distribution (akin to those originally found by Herschel in the distribution of stars). The distribution of nearby galaxies do indeed show such gaps. However, if these gaps are due to clouds of intergalactic dust one ought to find that they are also devoid of more distant galaxies. This is not the case, which is contrary to what one would expect if the gaps were due to clouds of intergalactic dust. One problem with the above described method to search for IGM dust is that on very large scales the three-dimensional distribution of clusters of galaxies is far from uniform and consists of large so-called *voids* with clusters draped around these voids. Clearly this means of searching for intergalacitic dust is not likely to be fruitful.

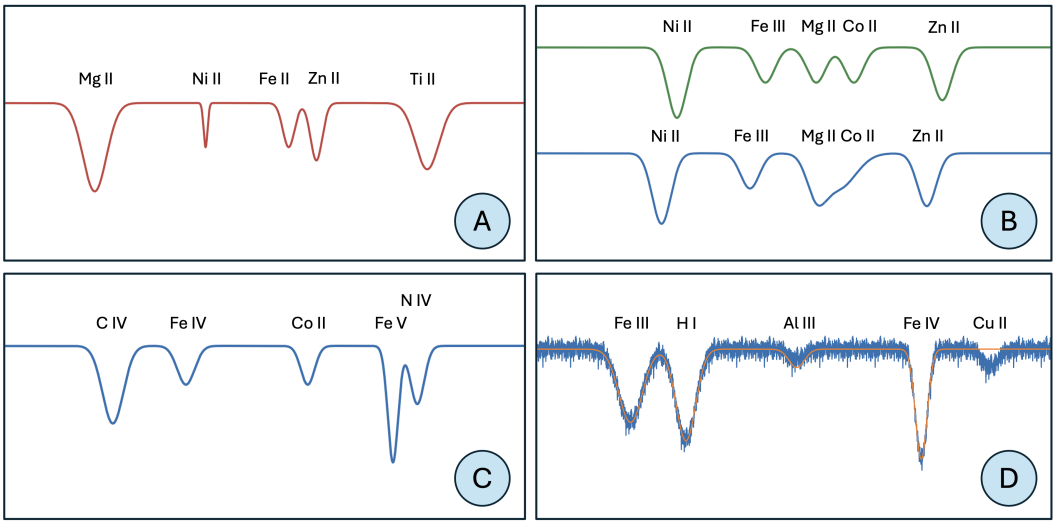


Figure 19.9: *Panels showing a hypothetical spectral window containing stellar and one interstellar line. The spectral resolution of the spectrum in panel A is very high; that of the other panels is so low that the line broadening is dominated by the spectral resolution of the instrument. Hints: In panel A the interstellar cloud in front of the star is cold relative to the stellar temperature. In panel B two observations of the star are shown taken several days apart. In panel D the noisy stellar spectrum is fitted with a model atmosphere.*

Exercise 19.1

In Fig. 19.9, each panel shows a hypothetical spectral window containing stellar and one interstellar line. Identify in each panel the interstellar line and briefly explain your reasoning. The spectral resolution of the spectrum in panel A is very high; that of the other panels is so low that the line broadening is dominated by the spectral resolution of the instrument. Hints: In panel A the interstellar cloud in front of the star is cold relative to the stellar temperature. In panel B two observations of the star are shown taken several days apart. In panel D the noisy stellar spectrum is fitted with a model atmosphere.

Exercise 19.2

The total mass of neutral gas in the Galaxy is $M_{\rm nISM} \sim 4 \times 10^9 M_\odot$. Assume that it is uniformly distributed in a disk of radius $R_{\rm disk} = 15\,$kpc and thickness $H = 200\,$pc, and that it is a mixture of H and He with $n_{\rm He}/n_{\rm H} = 0.1$. We further assume that 0.7% of the interstellar mass is in the form of dust in spherical particles of radius $a = 1000\,\mathring{A} = 0.1\,\mu$m and material density $\rho_{\rm s} = 2\,{\rm gr\,cm^{-3}}$. For such a grain, let $Q_{\rm ext}$ be the ratio of the visual (V band, $\lambda = 0.55\,\mu$m) extinction cross section (see Eq. 4.5) to the geometric cross section $\pi\, a^2$. Suppose that $Q_{\rm ext} \approx 1$.

a) Compute the average number density of hydrogen nuclei within the disk.

b) Compute the average number density of dust grains in interstellar space.

c) Compute the visual extinction A_V in magnitudes between the Sun and the Galactic Center (assumed to be $L = 8.5$ kpc away).

In reality, the interstellar medium is not homogeneous but interspersed with gas clouds. We assume that 30% of the gas and dust mass is in spherical molecular clouds of radius 15 pc and mean density $n(H_2) = 100\,\mathrm{cm}^{-3}$.

d) What would be the mass of one such cloud? How many such molecular clouds would there be in the Galaxy?

e) We consider the clouds to be distributed randomly according to a Poisson random point field, i.e. each point (or cloud in our case) is stochastically independent to all the other points in the process. The Poisson point process is related to the Poisson distribution, which implies that the probability of a Poisson random variable N being equal to n is given by

$$P\{N = n\} \quad = \quad \frac{\Lambda^n}{n!}\, e^{-\Lambda}, \tag{19.9}$$

where Λ defines the properties of the Poisson distribution. In our case, Λ is the expectation value for the number of clouds that will be intersected by the line of sight to the Galactic center. Compute Λ.

f) Compute the probability that our line of sight to the Galactic center happens *not* to intersect any molecular clouds. Assuming hydrogen and dust in the inter-cloud medium to be distributed uniformly throughout the disk volume, what will be the visual extinction to the Galactic center?

Exercise 19.3

For the Cardelli interstellar extinction law the constant the color excess ratio $E(U-B)/E(B-V)$ is slightly different from the value 0.72 given in Eq. (19.7). Use Table 19.3 (where $R_V = 3.1$ is adopted) to compute this value for the Cardelli law.

Exercise 19.4

a) Show, using Eq. (19.7), that indeed the quantity Q in Eq. (19.8) is independent of reddening.

b) The extinction vector drawn in Fig. 19.5 is drawn such that, adopting $R_V = 3.1$, for an extinction $A_V = 1$, the observed $U-B$ reddens by 0.23 and the observed $B-V$ by 0.32. Verify that this is indeed what is to be expected.

Appendix

List of Tables

LumType / SpType	Criteria		
	V	III	I
O2-O3	He I + II λ4026 < He II λ4200; weak N III 4634-40-42 emission[a] ↑[b]; weak Si IV 4089-4116 emission ↑; median weak N IV 4058 emission ↑; weak N V 4604-20 ↓		
	weak He II 4686	weaker He II 4686 than class V	weak He II 4686 emission
O4	He I + II λ4026 < He II λ4200; weak N III 4634-40-42 emission ↑; weak Si IV 4089-4116 emission ↑; weak N V 4604-20 ↓		
	weak He II 4686 weak He I 4471	weaker He II 4686 than class V	weak He II 4686 emission; N IV 4058 emission
O5	He II λ4541 > He I λ4471; N III 4634-40-42 emission ↑		
	He II λ4686 ≈ He II λ4541 He II 4686	He II λ4686 < He II λ4541 weak He II 4686 weak Si IV 4116 emission	N III λ4634-40-42 ≈ He II λ4686 He II 4686 emission Si IV 4116 emission
O6	He I + II λ4026 ≈ He II λ4200; N III 4634-40-42 emission ↑; Si IV 4089-4116 ↑		
	He II λ4541 > He I λ4471 median weak He II 4686	He II λ4541 ≈ He I λ4471 weak He II 4686	He II λ4541 < He I λ4471 He II 4686 emission; weak N III λ4097
O7	He I λ4471 ≈ He II λ4541; N III 4634-40-42 emission ↑; weak Si IV 4089-4116 ↑		
	He II λ4686 > He I λ4471	He II λ4686 ≈ He I λ4471	Si IV 4686-4504 emission He II 4686 emission
O8-O9	He I λ4471 > He II λ4541; weak Si IV 4089 ↑		
	C III λ4650 > He II λ4686 He II λ4686 > He I λ4713 Si IV λ4089 < C III λ4187	C III λ4650 > He II λ4686 He II λ4686 > He I λ4713 Si IV λ4089 ≈ C III λ4187 weak Si IV 4686-4504 emission	Si IV 4686-4504 emission He II 4686 emission weak N III 4097 N III 4634-40-42 emission
B0	Si III 4552-68-75; weak He II 4200; O II 4070-76 ↑; Si IV 4089 ↑; He II 4686↓		
	He II λ4686 > He I λ4711 weak O II 4350	He II λ4686 < He I λ4713 O II 4350, N II 3995	Si IV λ4089 > He I λ4121 strong O II 4350; N III 4097; N II 3995
B1	Si III 4552-68-75 ↑; N II 3995 ↑; O II 4070-76 ↑; Si IV 4089 ↑		
	Si III λ4552 ≈ Si III λ4568 Si III λ4568 ≈ Si III λ4575	Si III λ4552 ≈ Si III λ4568 Si III λ4568 ≈ Si III λ4575 weak O II 4350	Si III λ4552 > Si III λ4568 Si III λ4568 > Si III λ4575 strong O II 4350
B2-B3	Si I λ4128-30< He I λ4121; He I λ4471 > Mg II λ4481; weak Si III 4552-68-75 ↑; C II 4267; O II 4070-76 ↑		
	weak He I 4121	N II 3995	weak O II 4350, N II 3995
B5	He I λ4471 > Mg II λ4481; C II 4267; Si II 4128-30		
	Si III 4552-68-75 absent	weak Si III 4552-68-75 weak Si III 4552-68-75	Si I λ4128-30< He I λ4121 median weak Si III 4552-68-75; N II 3995
B7	He I λ4471 > Mg II λ4481; C II 4267; weak Fe II 4233		
	He I 4121 absent	Si I λ4128-30≈ He I λ4121	Si I λ4128-30< He I λ4121; weak N II 3995
B8	weak Fe II 4233; weak Si II 4128-30 ↑		
	He I λ4471 < Mg II λ4481	He I λ4471 < Mg II λ4481	He I λ4471 ≈ Mg II λ4481 Si I λ4128-30> He I λ4144; C II 4267
B9	He I λ4471 < Mg II λ4481; weak Si II 4128-30 ↑; weak Fe II 4233		
	weak He I 4026	Si II λ4128-30≈ He I λ4026	Si II λ4128-30> He I λ4026

Table B.1: *Criteria for classifying sub-spectral type and luminosity class of O and B stars.* [a] *Represents emission lines; all others are absorption lines.* [b] *Means the lines strengthen with luminosity class from V to I. From: Liu et al. (2019).*

SpT	$T_{\rm eff}$	$\log L$	$M_{\rm bol}$	$\rm BC_V$	M_V	$B-V$	G-V	Bp-Rp	G-Rp	$\rm M_G$	$U-B$	V-Rc	V-Ic	V-Ks	J-H	H-Ks	$\rm M_J$	$\rm M_{Ks}$	i-z	z-Y	$R_\odot$	$M_\odot$
O3V	46000	5.80	-9.75	-4.05	-5.7	-0.32	...	...	...	...	-1.22	...	...	...	...	...	...	...	...	...	12.5	...
O4V	43000	5.67	-9.42	-3.92	-5.5	-0.32	...	...	...	...	-1.20	...	...	...	...	...	...	...	...	...	12.3	. ..
O5V	41500	5.58	-9.20	-3.77	-5.4	-0.32	...	...	...	...	-1.19	...	...	...	...	...	...	...	...	...	11.9	...
O5.5V	40000	5.46	-8.90	-3.65	-5.2	-0.32	...	...	...	...	-1.18	...	...	...	...	...	...	...	...	...	11.2	...
O6V	39000	5.37	-8.67	-3.57	-5.1	-0.32	...	...	...	...	-1.17	...	...	...	...	...	...	...	...	...	10.6	...
O6.5V	37300	5.24	-8.35	-3.45	-4.9	-0.32	...	...	...	...	-1.16	...	...	...	...	...	...	...	...	...	10.0	...
O7V	36500	5.17	-8.18	-3.38	-4.8	-0.32	...	...	...	...	-1.15	...	...	...	...	...	...	...	...	...	9.62	28
O7.5V	35000	5.05	-7.87	-3.27	-4.6	-0.32	...	...	...	...	-1.14	...	...	...	...	...	...	...	...	...	9.11	24
O8V	34500	4.99	-7.72	-3.22	-4.5	-0.32	...	...	...	...	-1.13	...	...	...	...	...	...	...	...	...	8.75	22.9
O8.5V	33000	4.87	-7.43	-3.13	-4.3	-0.32	...	...	...	...	-1.12	...	...	...	...	...	...	...	...	...	8.33	20.5
O9V	32500	4.82	-7.29	-3.09	-4.2	-0.318	...	...	...	...	-1.114	...	-0.369	-1.000	-0.164	-0.071	-3.44	-3.20	...	...	8.11	19.7
O9.5V	32000	4.76	-7.15	-3.06	-4.1	-0.312	...	...	...	...	-1.087	...	-0.361	-0.977	-0.161	-0.069	-3.35	-3.12	...	...	7.80	18.5
B0V	31500	4.70	-7.02	-3.02	-4.0	-0.307	...	...	...	...	-1.067	...	-0.355	-0.958	-0.159	-0.067	-3.27	-3.04	...	...	7.53	17.5
B0.5V	29000	4.47	-6.43	-2.87	-3.6	-0.295	...	...	...	...	-1.026	...	-0.338	-0.913	-0.153	-0.063	-2.90	-2.69	...	...	6.81	15
B1V	26000	4.13	-5.68	-2.61	-3.1	-0.278	...	...	...	...	-0.995	-0.115	-0.325	-0.874	-0.148	-0.059	-2.43	-2.23	...	...	5.72	11
B1.5V	24500	3.89	-5.24	-2.43	-2.8	-0.252	...	...	...	...	-0.910	-0.114	-0.281	-0.752	-0.132	-0.047	-2.23	-2.05	...	...	4.89	10
B2V	20600	3.38	-3.71	-2.06	-1.7	-0.210	...	...	...	...	-0.790	-0.094	-0.230	-0.602	-0.113	-0.032	-1.24	-1.10	...	...	3.85	7.3
B2.5V	18500	3.10	-3.15	-1.79	-1.4	-0.198	...	...	...	...	-0.732	-0.087	-0.210	-0.544	-0.105	-0.026	-0.99	-0.86	...	...	3.45	6.1
B3V	17000	2.96	-2.64	-1.58	-1.1	-0.178	...	...	...	...	-0.673	-0.080	-0.192	-0.492	-0.098	-0.021	-0.73	-0.61	...	...	3.48	5.4
B4V	16700	2.91	-2.52	-1.53	-1.0	-0.165	...	...	...	...	-0.619	-0.074	-0.176	-0.447	-0.092	-0.016	-0.66	-0.55	...	...	3.41	5.0
B5V	15700	2.80	-2.25	-1.35	-0.9	-0.156	...	...	...	...	-0.581	-0.070	-0.165	-0.417	-0.089	-0.013	-0.59	-0.48	...	...	3.40	4.6
B6V	14500	2.54	-1.66	-1.16	-0.5	-0.140	...	...	...	...	-0.504	-0.062	-0.145	-0.358	-0.081	-0.007	-0.23	-0.14	...	...	2.95	4.0
B7V	14000	2.49	-1.47	-1.07	-0.4	-0.128	...	...	...	...	-0.459	-0.058	-0.133	-0.325	-0.077	-0.004	-0.16	-0.08	...	...	2.99	3.9
B8V	12500	2.27	-0.91	-0.81	-0.2	-0.109	...	...	...	...	-0.364	-0.048	-0.108	-0.254	-0.067	0.003	-0.01	0.05	...	...	2.91	3.4
B9V	10700	1.79	0.28	-0.42	0.7	-0.070	0.017	-0.120	-0.036	0.72	-0.200	-0.028	-0.061	-0.121	-0.050	-0.054	0.79	0.82	...	...	2.28	2.8
B9.5V	10400	1.73	0.44	-0.38	0.8	-0.050	0.015	-0.087	-0.016	0.82	-0.130	-0.017	-0.035	-0.048	-0.044	-0.043	0.83	0.85	...	...	2.26	2.5
A0V	9700	1.54	0.87	-0.24	1.11	0.000	0.007	-0.037	0.010	1.09	-0.005	0.001	0.004	0.041	-0.032	-0.022	1.07	1.07	...	...	2.09	2.3
A1V	9200	1.41	1.19	-0.15	1.34	0.040	0.000	0.011	0.036	1.32	0.033	0.019	0.044	0.101	-0.024	-0.023	1.25	1.24	...	...	2.00	2.15
A2V	8840	1.33	1.38	-0.10	1.48	0.070	-0.005	0.068	0.069	1.46	0.063	0.042	0.091	0.188	-0.010	-0.029	1.32	1.29	...	...	1.97	2.05
A3V	8550	1.29	1.49	-0.06	1.55	0.090	-0.009	0.096	0.084	1.52	0.077	0.050	0.108	0.228	-0.002	-0.029	1.35	1.32	...	...	2.01	2.00
A4V	8270	1.20	1.72	-0.04	1.76	0.140	-0.020	0.166	0.122	1.73	0.097	0.078	0.164	0.353	0.022	-0.021	1.47	1.41	...	...	1.94	1.90
A5V	8080	1.16	1.81	-0.03	1.84	0.160	-0.024	0.194	0.137	1.81	0.100	0.089	0.186	0.403	0.031	-0.021	1.51	1.44	...	...	1.94	1.85
A6V	8000	1.14	1.87	-0.02	1.89	0.170	-0.026	0.208	0.145	1.86	0.098	0.094	0.197	0.428	0.036	-0.018	1.54	1.46	...	...	1.93	1.83
A7V	7800	1.06	2.07	0.00	2.07	0.210	-0.036	0.263	0.174	2.03	0.091	0.117	0.242	0.528	0.055	-0.017	1.64	1.54	...	...	1.86	1.76
A8V	7500	0.97	2.29	0.00	2.29	0.250	-0.046	0.320	0.204	2.25	0.082	0.140	0.288	0.626	0.075	-0.009	1.78	1.66	...	...	1.81	1.67
A9V	7440	0.97	2.30	0.00	2.30	0.255	-0.047	0.327	0.208	2.26	0.080	0.143	0.294	0.638	0.078	-0.007	1.78	1.66	...	...	1.84	1.67
F0V	7220	0.89	2.50	-0.01	2.51	0.290	-0.060	0.377	0.230	2.46	0.053	0.166	0.339	0.732	0.098	0.003	1.90	1.76	...	...	1.79	1.59
F1V	7030	0.77	2.78	-0.01	2.79	0.330	-0.079	0.434	0.252	2.73	0.021	0.190	0.385	0.828	0.119	0.009	2.13	1.96	...	...	1.64	1.50
F2V	6810	0.70	2.97	-0.02	2.99	0.370	-0.093	0.490	0.279	2.92	-0.008	0.213	0.432	0.925	0.140	0.011	2.26	2.07	...	...	1.61	1.44
F3V	6720	0.67	3.05	-0.03	3.08	0.390	-0.100	0.518	0.293	3.01	-0.016	0.222	0.449	0.961	0.147	0.008	2.32	2.12	...	...	1.60	1.43
F4V	6640	0.61	3.19	-0.04	3.23	0.410	-0.107	0.546	0.307	3.16	-0.026	0.236	0.476	1.017	0.159	0.000	2.42	2.21	...	...	1.53	1.39
F5V	6510	0.54	3.36	-0.04	3.40	0.440	-0.116	0.587	0.329	3.32	-0.029	0.252	0.506	1.079	0.173	-0.004	2.55	2.32	...	...	1.46	1.33
F6V	6340	0.43	3.65	-0.05	3.70	0.484	-0.129	0.640	0.356	3.61	-0.021	0.276	0.553	1.185	0.199	-0.012	2.78	2.52	...	...	1.36	1.25
F7V	6240	0.36	3.81	-0.06	3.87	0.510	-0.135	0.670	0.372	3.77	-0.012	0.290	0.579	1.244	0.213	-0.013	2.90	2.63	...	...	1.30	1.21
F8V	6170	0.31	3.96	-0.07	4.01	0.530	-0.140	0.694	0.385	3.90	0.001	0.300	0.599	1.290	0.225	-0.016	3.01	2.72	...	...	1.25	1.18
F9V	6060	0.26	4.07	-0.08	4.15	0.552	-0.146	0.719	0.399	4.04	0.014	0.312	0.620	1.340	0.237	-0.014	3.11	2.81	...	...	1.23	1.14
F9.5V	6000	0.21	4.22	-0.08	4.29	0.572	-0.150	0.767	0.431	4.17	0.033	0.323	0.640	1.385	0.249	-0.012	3.22	2.91	...	...	1.18	1.11

Table B.2: *Dwarf stellar, brown dwarf, and exo-planet color and effective temperature sequence from Pecaut & Mamajek (2013, ApJS 208, 9). The full table is available on line and kept up to date. Colors are based on Johnson UBV, Tycho BtVt, Cousins RcIc, Gaia DR2 G/Bp/Rp, Sloan Iz, 2MASS JHSs. G-V color is Gaia DR2 G - Johnson V. The original table also provides WISE colors, which are left out here to keep the table within page size.* *Table continues on the next page.*

SpT	$T_{\rm eff}$	$\log L$	$M_{\rm bol}$	BC$_V$	M_V	$B-V$	G-V	Bp-Rp	G-Rp	M$_G$	$U-B$	V-Rc	V-Ic	V-Ks	J-H	H-Ks	M$_J$	M$_{Ks}$	i-z	z-Y	$R_\odot$	$M_\odot$
G0V	5920	0.14	4.40	-0.09	4.45	0.596	-0.155	0.782	0.438	4.32	0.058	0.336	0.664	1.440	0.262	0.067	3.34	3.01	...	...	1.12	1.08
G1V	5880	0.13	4.40	-0.10	4.50	0.625	-0.162	0.803	0.448	4.37	0.067	0.340	0.672	1.458	0.267	0.068	3.38	3.04	...	...	1.12	1.07
G2V	5770	0.01	4.68	-0.11	4.79	0.650	-0.167	0.823	0.459	4.65	0.133	0.363	0.713	1.564	0.293	0.073	3.57	3.20	...	...	1.01	1.02
G3V	5720	-0.01	4.74	-0.12	4.86	0.660	-0.169	0.832	0.464	4.72	0.152	0.368	0.722	1.590	0.299	0.074	3.64	3.27	...	...	1.01	1.00
G4V	5680	-0.04	4.80	-0.13	4.94	0.670	-0.172	0.841	0.468	4.79	0.175	0.374	0.733	1.621	0.307	0.075	3.70	3.32	...	...	0.986	0.99
G5V	5660	-0.05	4.84	-0.13	4.98	0.680	-0.174	0.850	0.473	4.83	0.185	0.377	0.738	1.635	0.310	0.076	3.74	3.35	...	...	0.982	0.98
G6V	5590	-0.11	4.98	-0.15	5.13	0.700	-0.180	0.869	0.483	4.97	0.227	0.388	0.758	1.691	0.324	0.079	3.84	3.44	...	...	0.939	0.97
G7V	5530	-0.12	5.02	-0.16	5.18	0.710	-0.182	0.880	0.489	5.02	0.243	0.393	0.766	1.712	0.329	0.080	3.88	3.47	...	...	0.949	0.96
G8V	5490	-0.17	5.15	-0.17	5.32	0.730	-0.188	0.900	0.499	5.15	0.284	0.404	0.786	1.768	0.342	0.082	3.97	3.55	...	...	0.909	0.94
G9V	5340	-0.25	5.34	-0.21	5.55	0.775	-0.204	0.950	0.524	5.34	0.358	0.423	0.820	1.861	0.365	0.087	4.14	3.69	...	...	0.876	0.90
K0V	5280	-0.33	5.54	-0.22	5.76	0.816	-0.221	0.98	0.56	5.53	0.436	0.443	0.853	1.953	0.387	0.091	4.29	3.81	...	...	0.817	0.87
K0.5V	5240	-0.33	5.57	-0.23	5.80	0.825	-0.225	1.00	0.56	5.56	0.456	0.448	0.862	1.977	0.393	0.092	4.31	3.82	...	...	0.828	0.86
K1V	5170	-0.36	5.65	-0.26	5.89	0.842	-0.232	1.01	0.56	5.63	0.491	0.457	0.879	2.021	0.402	0.094	4.37	3.87	...	...	0.814	0.85
K1.5V	5140	-0.39	5.70	-0.27	5.97	0.860	-0.241	1.05	0.58	5.69	0.528	0.467	0.895	2.066	0.412	0.096	4.41	3.90	...	...	0.809	0.82
K2V	5040	-0.47	5.88	-0.29	6.19	0.884	-0.254	1.10	0.62	5.89	0.581	0.482	0.920	2.132	0.427	0.098	4.57	4.04	...	...	0.763	0.78
K2.5V	4990	-0.51	6.02	-0.32	6.34	0.938	-0.288	1.17	0.64	6.04	0.691	0.513	0.974	2.274	0.459	0.104	4.63	4.07	...	...	0.742	0.76
K3V	4830	-0.58	6.16	-0.41	6.57	0.986	-0.322	1.21	0.66	6.26	0.776	0.537	1.013	2.380	0.483	0.109	4.76	4.16	...	...	0.729	0.75
K3.5V	4700	-0.64	6.34	-0.45	6.79	1.050	-0.371	1.27	0.68	6.42	0.917	0.592	1.108	2.582	0.520	0.118	4.85	4.21	...	...	0.720	0.73
K4V	4600	-0.67	6.42	-0.56	6.98	1.100	-0.412	1.32	0.70	6.55	1.004	0.640	1.190	2.733	0.544	0.125	4.92	4.25	...	...	0.726	0.72
K4.5V	4540	-0.68	6.44	-0.60	7.04	1.116	-0.425	1.38	0.72	6.62	1.028	0.654	1.216	2.781	0.552	0.127	4.94	4.26	...	...	0.737	0.71
K5V	4410	-0.78	6.68	-0.68	7.36	1.150	-0.454	1.45	0.74	6.95	1.081	0.685	1.272	2.883	0.568	0.132	5.18	4.48	...	...	0.698	0.68
K5.5V	4330	-0.84	6.85	-0.75	7.60	1.200	-0.495	1.51	0.76	7.16	1.144	0.728	1.357	3.034	0.591	0.139	5.30	4.57	...	...	0.672	0.66
K6V	4230	-0.90	6.99	-0.81	7.80	1.240	-0.528	1.58	0.79	7.32	1.184	0.759	1.420	3.143	0.601	0.148	5.41	4.66	...	...	0.661	0.65
K6.5V	4190	-0.96	7.09	-0.92	8.01	1.310	-0.581	1.64	0.83	7.49	1.213	0.796	1.505	3.288	0.613	0.159	5.56	4.78	...	...	0.656	0.64
K7V	4070	-0.98	7.18	-0.95	8.15	1.330	-0.595	1.71	0.87	7.58	1.221	0.806	1.529	3.330	0.617	0.162	5.60	4.82	...	...	0.654	0.63
K8V	4000	-1.10	7.44	-1.03	8.47	1.380	-0.628	1.72	0.88	7.84	1.216	0.843	1.632	3.487	0.623	0.176	5.78	4.98	...	...	0.587	0.59
K9V	3940	-1.18	7.59	-1.10	8.69	1.420	-0.69	1.78	0.90	8.00	1.210	0.866	1.699	3.584	0.625	0.184	5.92	5.11	...	...	0.552	0.56
M0V	3870	-1.20	7.75	-1.16	8.91	1.408	-0.65	1.84	0.92	8.26	1.204	0.889	1.766	3.680	0.626	0.193	6.04	5.22	0.33	...	0.559	0.55
M0.5V	3800	-1.27	7.90	-1.38	9.20	1.441	-0.74	1.97	0.96	8.46	1.184	0.924	1.886	3.84	0.620	0.208	6.19	5.36	0.37	...	0.535	0.54
M1V	3700	-1.40	8.25	-1.44	9.69	1.475	-0.82	2.09	1.01	8.87	1.172	0.959	2.019	4.02	0.613	0.225	6.51	5.67	0.41	...	0.496	0.49
M1.5V	3650	-1.47	8.40	-1.57	9.97	1.486	-0.85	2.14	1.03	9.12	1.170	0.978	2.089	4.12	0.607	0.228	6.69	5.85	0.47	...	0.460	0.47
M2V	3550	-1.57	8.65	-1.65	10.30	1.500	-0.92	2.25	1.06	9.38	1.170	1.001	2.173	4.24	0.600	0.234	6.89	6.06	0.53	...	0.434	0.44
M2.5V	3500	-1.68	8.94	-1.76	10.70	1.522	-1.02	2.39	1.10	9.68	1.175	1.041	2.306	4.43	0.589	0.244	7.01	6.27	0.57	...	0.393	0.40
M3V	3410	-1.78	9.17	-1.97	11.14	1.544	-1.09	2.49	1.12	10.05	1.181	1.079	2.420	4.60	0.579	0.252	7.40	6.54	0.61	...	0.369	0.36
M3.5V	3250	-2.07	9.92	-2.27	12.19	1.602	-1.29	2.71	1.17	10.90	1.200	1.178	2.680	5.00	0.558	0.269	8.02	7.19	0.66	...	0.291	0.26
M4V	3200	-2.20	10.21	-2.59	12.80	1.661	-1.41	2.95	1.24	11.39	1.222	1.241	2.831	5.25	0.557	0.282	8.39	7.55	0.71	...	0.258	0.22
M4.5V	3100	-2.31	10.52	-3.05	13.57	1.72	-1.55	3.13	1.27	12.02	1.23	1.345	3.073	5.64	0.564	0.301	8.79	7.93	0.81	...	0.243	0.18
M5V	3030	-2.52	11.02	-3.28	14.30	1.874	-1.74	3.36	1.33	12.56	1.24	1.446	3.277	5.94	0.580	0.311	9.25	8.36	0.91	0.47	0.199	0.16
M5.5V	3000	-2.79	11.71	-3.80	15.51	1.91	-2.01	3.70	1.37	13.50	1.3	1.656	3.664	6.50	0.588	0.329	9.93	9.01	1.13	0.52	0.149	0.12
M6V	2850	-3.02	12.26	-4.36	16.62	2.00	-2.14	3.95	1.45	14.48	1.3	1.950	4.13	7.30	0.605	0.352	10.28	9.32	1.45	0.60	0.127	0.10
M6.5V	2710	-3.09	12.47	-4.60	17.07	2.06	-2.64	4.41	1.48	14.32	...	2.003	4.31	7.60	0.609	0.364	10.47	9.47	1.58	0.64	0.129	0.097
M7V	2650	-3.21	12.75	-5.06	17.81	2.06	-2.98	4.75	1.52	14.83	...	2.180	4.45	8.05	0.613	0.386	10.76	9.76	1.77	0.70	0.118	0.090
M7.5V	2600	-3.29	12.97	-5.46	18.42	2.17	-3.03	4.77	1.54	15.33	...	2.160	4.56	8.45	0.650	0.422	10.68	9.97	1.85	0.74	0.112	0.089
M8V	2500	-3.36	13.14	-5.70	18.84	2.20	-3.08	4.80	1.57	15.73	...	2.150	4.64	8.73	0.670	0.450	11.23	10.11	1.93	0.77	0.111	0.082
M8.5V	2440	-3.44	13.34	-5.80	19.14	...	-3.04	5.03	1.58	16.05	...	1.967	4.71	8.92	0.685	0.470	11.45	10.22	1.96	0.80	0.107	0.081
M9V	2400	-3.57	13.67	-5.90	19.36	...	-3.00	4.78	1.59	16.29	...	1.890	4.75	9.00	0.749	0.480	11.53	10.30	1.99	0.82	0.095	0.079
M9.5V	2320	-3.55	13.62	-6.13	19.75	...	-3.10	...	1.63	16.33	...	2.510	4.79	9.30	0.826	0.505	11.78	10.45	2.00	0.84	0.104	0.078
L0V	2250	-3.57	13.67	-6.33	20.0	...	-3.20	...	1.68	16.36	...	...	4.82	9.45	0.79	0.50	11.84	10.55	2.01	0.86	0.108	0.077
L1V	2100	-3.70	13.99	-6.51	20.5	...	...	...	1.66	16.83	...	...	4.91	9.70	0.80	0.54	12.14	10.80	2.02	0.88	0.107	0.076
L2V	1960	-3.84	14.34	-6.56	20.9	...	...	...	1.69	17.24	...	...	5.05	10.0	0.87	0.57	12.34	10.90	2.04	0.90	0.104	0.075
L3V	1830	-3.98	14.96	-6.74	21.7	...	...	...	1.70	17.76	...	...	5.29	10.4	1.00	0.63	12.93	11.30	2.10	0.92	0.102	...
L4V	1700	-4.11	15.02	-7.28	22.3	...	...	...	1.74	18.32	...	...	5.57	10.9	1.14	0.63	13.17	11.40	2.20	0.94	0.102	...
L5V	1590	-4.24	15.34	-7.76	23.1	...	...	...	1.71	18.86	...	...	6.28	11.4	1.13	0.65	13.60	11.82	2.33	0.97	0.100	...

Table B.3: *Table continued from previous page.*

SpT	$T_{\rm eff}$	$\log L$	$M_{\rm bol}$	BC$_V$	M_V	$B-V$	G-V	Bp-Rp	G-Rp	M$_G$	$U-B$	V-Rc	V-Ic	V-Ks	J-H	H-Ks	M$_J$	M$_{Ks}$	i-z	z-Y	$R_\odot$	$M_\odot$
L6V	1490	-4.36	15.64	...	...	...	...	...	1.74	19.25	...	...	...	...	1.08	0.64	13.99	12.27	2.51	1.00	0.099	...
L7V	1410	-4.46	15.89	...	...	...	...	...	2.11	19.30	...	...	...	...	1.10	0.62	14.34	12.62	2.71	1.04	0.099	...
L8V	1350	-4.55	16.12	...	...	...	...	...	1.87	20.00	...	...	...	...	1.14	0.63	14.47	12.70	2.93	1.09	0.097	...
L9V	1300	-4.61	16.27	...	...	...	...	...	...	...	...	...	...	...	1.10	0.63	14.47	12.74	3.15	1.16	0.098	...
T0V	1260	-4.66	16.39	...	...	...	...	...	...	...	...	...	...	...	1.02	0.54	14.46	12.90	3.36	1.23	0.098	...
T1V	1230	-4.69	16.47	...	...	...	...	...	...	...	...	...	...	...	1.02	0.45	14.34	12.87	3.55	1.33	0.100	...
T2V	1200	-4.73	16.57	...	...	...	...	...	...	...	...	...	...	...	0.86	0.27	14.32	13.19	3.70	1.43	0.100	...
T3V	1160	-4.77	16.67	...	...	...	...	...	...	...	...	...	...	...	0.68	0.08	14.45	13.69	3.82	1.55	0.102	...
T4V	1120	-4.84	16.84	...	...	...	...	...	...	...	...	...	...	...	0.35	-0.19	14.67	14.51	3.90	1.68	0.101	...
T4.5V	1090	-4.90	16.99	...	...	...	...	...	...	...	...	...	...	...	0.2	-0.06	14.80	14.66	3.93	1.75	0.099	...
T5V	1050	-4.95	17.12	...	...	...	...	...	...	...	...	...	...	...	0.2	-0.08	15.02	14.90	3.95	1.81	0.101	...
T5.5V	1010	-5.04	17.34	...	...	...	...	...	...	...	...	...	...	...	0.2	-0.10	15.28	15.18	3.97	1.89	0.099	...
T6V	960	-5.12	17.54	...	...	...	...	...	...	...	...	...	...	...	0.1	-0.03	15.61	15.54	3.98	1.96	0.100	...
T7V	840	-5.37	18.17	...	...	...	...	...	...	...	...	...	...	...	0.0	0.00	15.77	16.36	4.01	2.11	0.098	...
T7.5V	770	-5.54	18.59	...	...	...	...	...	...	...	...	...	...	...	0.2	-0.05	16.51	16.85	4.05	2.19	0.095	...
T8V	700	-5.71	19.02	...	...	...	...	...	...	...	...	...	...	...	0.2	-0.05	17.24	17.43	4.08	2.26	0.095	...
T8.5V	610	-5.93	19.57	...	...	...	...	...	...	...	...	...	...	...	0.2	...	18.45	...	...	...	0.097	...
T9V	530	-6.15	20.12	...	...	...	...	...	...	...	...	...	...	...	0.1	-0.20	19.67	18.48	...	...	0.100	...
T9.5V	475	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	19.52:	19.32	...	...	...	...
Y0V	420	-6.52	...	...	...	...	...	...	...	...	...	...	...	...	...	-0.5	21	...	...	...	0.104	...
Y0.5V	390	...	...	...	...	...	...	...	...	...	...	...	...	...	...	-0.6	22	21.5	...	...	...	...
Y1V	350	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	23	23	...	...	...	...
Y1.5V	325	...	...	...	...	...	...	...	...	...	...	...	...	...	...	-0.8	25	23.5	...	...	...	...
Y2V	250	...	26.2	...	...	...	...	...	...	...	...	...	...	...	...	...	28.3	24	...	...	...	...

Table B.4: *Table continued from previous page.*

Sp	M_V	$B-V$	$U-B$	$V-R$	$R-I$	$V-K$	$J-H$	$H-K$	$K-L$	$T_{\rm eff}$	BC	$\log L$
MAIN SEQUENCE, V												
O5	-5.7	-0.33	-1.19	-0.15	-0.32					42 000	-4.40	5.940
O9	-4.5	-0.31	-1.12	-0.15	-0.32	-0.87	-0.14	-0.04	-0.06	34 000	-3.33	5.032
B0	-4.0	-0.30	-1.08	-0.13	-0.29	-0.83	-0.12	-0.04	-0.06	30 000	-3.16	4.764
B2	-2.45	-0.24	-0.84	-0.10	-0.22	-0.66	-0.09	-0.03	-0.05	20 900	-2.35	3.820
B5	-1.2	-0.17	-0.58	-0.06	-0.16	-0.42	-0.06	-0.01	-0.04	15 200	-1.46	2.964
B8	-0.25	-0.11	-0.34	-0.02	-0.10	-0.24	-0.03	0.00	-0.04	11 400	-0.80	2.320
A0	+0.65	-0.02	-0.02	0.02	-0.02	0.00	0.00	0.00	0.00	9 790	-0.30	1.760
A2	+1.3	+0.05	+0.05	0.08	0.01	0.14	0.02	0.01	0.01	9 000	-0.20	1.460
A5	+1.95	+0.15	+0.10	0.16	0.06	0.38	0.06	0.02	0.02	8 180	-0.15	1.180
F0	+2.7	+0.30	+0.03	0.30	0.17	0.70	0.13	0.03	0.03	7 300	-0.09	0.856
F2	+3.6	+0.35	+0.00	0.35	0.20	0.82	0.17	0.04	0.03	7 000	-0.11	0.504
F5	+3.5	+0.44	-0.02	0.40	0.24	1.10	0.23	0.04	0.04	6 650	-0.14	0.556
F8	+4.0	+0.52	+0.02	0.47	0.29			0.05		6 250	-0.16	0.364
G0	+4.4	+0.58	+0.06	0.50	0.31	1.41	0.31	0.05	0.05	5 940	-0.18	0.212
G2	+4.7	+0.63	+0.12	0.53	0.33	1.46	0.32	0.05	0.05	5 790	-0.20	0.100
G5	+5.1	+0.68	+0.20	0.54	0.35			0.06	0.05	5 560	-0.21	-0.056
G8	+5.5	+0.74	+0.30	0.58	0.38					5 310	-0.40	-0.140
K0	+5.9	+0.81	+0.45	0.64	0.42	1.96	0.45	0.08	0.06	5 150	-0.31	-0.336
K2	+6.4	+0.91	+0.64	0.74	0.48	2.22	0.50	0.09	0.07	4 830	-0.42	-0.492
K5	+7.35	+1.15	+1.08	0.99	0.63	2.85	0.61	0.11	0.10	4 410	-0.72	-0.752
M0	+8.8	+1.40	+1.22	1.28	0.91	3.65	0.67	0.17	0.14	3 840	-1.38	-1.068
M2	+9.9	+1.49	+1.18	1.50	1.19	4.11	0.66	0.20	0.16	3 520	-1.89	-1.304
M5	+12.3	+1.64	+1.24	1.80	1.67	6.17	0.62	0.33	0.29	3 170	-2.73	-1.928
GIANTS, III												
G5	+0.9	+0.86	+0.56	0.69	0.48					5 050	-0.34	1.676
G8	+0.8	+0.94	+0.70	0.70	0.48	2.16	0.50	0.09	0.06	4 800	-0.42	1.748
K0	+0.7	+1.00	+0.84	0.77	0.53	2.31	0.54	0.10	0.07	4 660	-0.50	1.820
K2	+0.5	+1.16	+1.16	0.84	0.58	2.70	0.63	0.12	0.09	4 390	-0.61	1.944
K5	-0.2	+1.50	+1.81	1.20	0.90	3.60	0.79	0.17	0.12	4 050	-1.02	2.388
M0	-0.4	+1.56	+1.87	1.23	0.94	3.85	0.83	0.19	0.12	3 690	-1.25	2.560
M2	-0.6	+1.60	+1.89	1.34	1.10	4.30	0.87	0.22	0.15	3 540	-1.62	2.788
M5	-0.3	+1.63	+1.58	2.18	1.96	5.96	0.95	0.29	0.20	3 380	-2.48	3.012
SUPERGIANTS, I												
O9	-6.5	-0.27	-1.13	-0.15	-0.32	-0.82	-0.05	-0.13	-0.08	32 000	-3.18	5.772
B2	-6.4	-0.17	-0.93	-0.05	-0.15	-0.40	-0.04	0.00	-0.07	17 600	-1.58	5.092
B5	-6.2	-0.10	-0.72	0.02	-0.07	-0.13	0.01	0.00	0.02	13 600	-0.95	4.760
B8	-6.2	-0.03	-0.55	0.02	0.00	0.07	0.07	-0.02	0.05	11 100	-0.66	4.644
A0	-6.3	-0.01	-0.38	0.03	0.05	0.19	0.09	-0.02	0.07	9 980	-0.41	4.584
A2	-6.5	+0.03	-0.25	0.07	0.07	0.32	0.12	-0.01	0.08	9 380	-0.28	4.612
A5	-6.6	+0.09	-0.08	0.12	0.13	0.48	0.13	0.02	0.07	8 610	-0.13	4.592
F0	-6.6	+0.17	+0.15	0.21	0.20	0.64	0.15	0.04	0.06	7 460	-0.01	4.544
F2	-6.6	+0.23	+0.18	0.26	0.21	0.75	0.18	0.05	0.06	7 030	-0.00	4.540
F5	-6.6	+0.32	+0.27	0.35	0.23	0.93	0.22	0.06	0.07	6 370	-0.03	4.552
F8	-6.5	+0.56	+0.41	0.45	0.27	1.21	0.28	0.07	0.07	5 750	-0.09	4.536
G0	-6.4	+0.76	+0.52	0.51	0.33	1.44	0.33	0.08	0.08	5 370	-0.15	4.520
G2	-6.3	+0.87	+0.63	0.58	0.40				0.08	5 190	-0.21	4.504
G5	-6.2	+1.02	+0.83	0.67	0.44					4 930	-0.33	4.512
G8	-6.1	+1.14	+1.07	0.69	0.46	1.99	0.43	0.11	0.09	4 700	-0.42	4.508
K0	-6.0	+1.25	+1.17	0.76	0.48	2.15	0.46	0.12	0.10	4 550	-0.50	4.500
K2	-5.9	+1.36	+1.32	0.85	0.55	2.43	0.52	0.13	0.12	4 310	-0.61	4.504
K5	-5.8	+1.60	+1.80	1.20	0.90	3.50	0.67	0.14	0.18	3 990	-1.01	4.624
M0	-5.6	+1.67	+1.90	1.23	0.94	3.80	0.73	0.18	0.20	3 620	-1.29	4.656
M2	-5.6	+1.71	+1.95	1.34	1.10	4.10	0.73	0.22	0.24	3 370	-1.62	4.788
M5	-5.6	+1.80	+1.60:	2.18	1.96					2 880	-3.47	5.528

Table B.5: *Calibration of the MK spectral type including their color indices. The colors are for the Johnson-Glass system, as described by Bessell & Brett (1988). Luminosity is in $L_\odot$; temperature is in K. For K5-M supergiants we have adopted a luminosity class Iab. Source: Cox (ed): Allen's Astrophysical Quantities, 4th edition, p. 388–389.*

Sp	Ia$^+$	Ia	Iab	Ib	II	III	IV	V
O3	4.642	4.650	4.667	4.673	4.660	4.639	4.618	4.699
O4	4.624	4.626	4.639	4.643	4.631	4.616	4.596	4.656
O4.5	4.614	4.614	4.626	4.629	4.617	4.606	4.588	4.638
O5	4.604	4.602	4.612	4.615	4.604	4.596	4.580	4.623
O5.5	4.593	4.590	4.599	4.601	4.592	4.587	4.574	4.609
O6	4.582	4.578	4.586	4.588	4.579	4.578	4.567	4.596
O6.5	4.570	4.565	4.573	4.574	4.567	4.570	4.562	4.585
O7	4.557	4.553	4.559	4.561	4.555	4.562	4.556	4.574
O7.5	4.544	4.540	4.546	4.548	4.543	4.553	4.550	4.564
O8	4.531	4.527	4.532	4.534	4.531	4.544	4.545	4.554
O8.5	4.517	4.513	4.519	4.521	4.518	4.536	4.539	4.545
O9	4.503	4.500	4.505	4.507	4.506	4.527	4.533	4.536
O9.5	4.457	4.457	4.462	4.465	4.467	4.497	4.511	4.508
O9.7	4.439	4.440	4.445	4.448	4.452	4.485	4.502	4.496
B0	4.410	4.414	4.419	4.422	4.427	4.465	4.486	4.477
B0.5	4.361	4.370	4.374	4.378	4.386	4.429	4.455	4.443
B1	4.312	4.326	4.330	4.334	4.343	4.389	4.420	4.405
B1.5	4.265	4.282	4.287	4.290	4.300	4.348	4.381	4.364
B2	4.219	4.241	4.245	4.247	4.257	4.305	4.339	4.320
B2.5	4.197	4.221	4.224	4.227	4.236	4.284	4.318	4.297
B3	4.175	4.201	4.204	4.207	4.216	4.262	4.296	4.274
B4	4.135	4.165	4.167	4.168	4.176	4.220	4.252	4.228
B5	4.098	4.131	4.132	4.132	4.139	4.179	4.208	4.182
B6	4.064	4.100	4.100	4.099	4.104	4.140	4.166	4.138
B7	4.033	4.072	4.071	4.068	4.072	4.103	4.125	4.096
B8	4.005	4.047	4.045	4.041	4.042	4.069	4.088	4.058
B9	3.980	4.024	4.021	4.016	4.016	4.038	4.053	4.023
A0	3.958	4.003	3.999	3.993	3.991	4.010	4.022	3.991
A1	3.944	3.990	3.985	3.978	3.976	3.992	4.002	3.973
A2	3.930	3.978	3.972	3.965	3.962	3.976	3.984	3.955
A4	3.905	3.954	3.947	3.939	3.936	3.946	3.952	3.926
A5	3.893	3.943	3.935	3.927	3.924	3.932	3.937	3.913
A6	3.880	3.931	3.923	3.915	3.912	3.919	3.924	3.901
A8	3.856	3.908	3.898	3.890	3.888	3.894	3.898	3.881
F0	3.831	3.883	3.873	3.865	3.864	3.869	3.875	3.864
F2	3.805	3.858	3.847	3.839	3.839	3.845	3.852	3.848
F4	3.779	3.831	3.819	3.811	3.813	3.820	3.828	3.832
F5	3.765	3.817	3.804	3.797	3.800	3.807	3.816	3.823
F6	3.751	3.802	3.790	3.782	3.786	3.794	3.804	3.815
F8	3.722	3.773	3.760	3.753	3.759	3.766	3.778	3.796
G0	3.694	3.744	3.730	3.723	3.730	3.738	3.751	3.774
G2	3.680	3.730	3.715	3.708	3.716	3.724	3.737	3.763
G4	3.666	3.715	3.700	3.694	3.702	3.710	3.723	3.751
G8	3.638	3.688	3.672	3.666	3.675	3.681	3.694	3.725
K0	3.625	3.675	3.658	3.652	3.663	3.668	3.679	3.712
K1	3.612	3.663	3.646	3.640	3.650	3.654	3.665	3.698
K3	3.586	3.639	3.622	3.617	3.628	3.629	3.637	3.671
K4	3.574	3.628	3.612	3.606	3.618	3.618	3.624	3.657
K5	3.561	3.618	3.601	3.597	3.609	3.607	3.611	3.644
K7	3.536	3.598	3.583	3.580	3.594	3.588	3.588	3.618
K9	3.509	3.579	3.567	3.566	3.582	3.573	3.568	3.595
M0	3.494	3.569	3.559	3.560	3.577	3.567	3.560	3.584
M1	3.461	3.549	3.543	3.548	3.569	3.556	3.545	3.564
M1.5	3.442	3.538	3.535	3.542	3.565	3.552	3.539	3.555
M2	3.422	3.526	3.527	3.536	3.561	3.549	3.533	3.547
M3	3.376	3.499	3.507	3.523	3.554	3.542	3.524	3.532
M4	3.322	3.467	3.485	3.507	3.546	3.536	3.516	3.517
M5	3.262	3.432	3.459	3.490	3.536	3.529	3.508	3.501
M6	3.200	3.395	3.433	3.471	3.526	3.522	3.498	3.482
M6.5	3.170	3.378	3.421	3.463	3.521	3.518	3.493	3.470
M7	3.143	3.363	3.411	3.456	3.517	3.514	3.487	3.457
M8	3.106	3.348	3.403	3.452	3.517	3.510	3.475	3.426
M9	3.109	3.365	3.422	3.472	3.534	3.515	3.465	3.388

Table B.6: *Temperature calibration of the MK spectral types for different luminosity classes. The logarithm of $T_{\rm eff}$ is given in Kelvin. Source: de Jager & Nieuwenhuijzen (1987, A&A 177, 217). In the article a fit-formula is given that provides the temperature as a function of a numerical spectral type and luminosity class. This formula (based on a 20th order Chebychev polynomial) is not used to generate the tabels provided here and in the original paper (that are based on a 40th order Chebychev polynomial).*

Sp	Ia$^+$	Ia	Iab	Ib	II	III	IV	V
O3	5.751	6.016	6.179	6.281	6.313	6.252	6.009	6.137
O4	5.822	5.997	6.114	6.177	6.159	6.047	5.807	5.904
O4.5	5.852	5.987	6.084	6.129	6.090	5.955	5.717	5.798
O5	5.880	5.977	6.055	6.084	6.025	5.870	5.635	5.698
O5.5	5.905	5.966	6.027	6.041	5.964	5.791	5.558	5.602
O6	5.927	5.955	5.999	6.000	5.905	5.715	5.485	5.510
O6.5	5.946	5.942	5.971	5.959	5.849	5.643	5.417	5.421
O7	5.962	5.929	5.943	5.919	5.794	5.574	5.351	5.335
O7.5	5.976	5.914	5.914	5.879	5.741	5.507	5.287	5.251
O8	5.987	5.898	5.886	5.840	5.689	5.442	5.225	5.168
O9.5	6.010	5.802	5.732	5.640	5.429	5.124	4.920	4.749
O9.7	6.009	5.775	5.693	5.590	5.366	5.046	4.845	4.661
B0	6.004	5.733	5.632	5.513	5.269	4.928	4.729	4.512
B0.5	5.986	5.658	5.526	5.381	5.103	4.723	4.527	4.258
B1	5.962	5.579	5.417	5.245	4.930	4.508	4.311	3.998
B1.5	5.934	5.499	5.306	5.106	4.752	4.282	4.081	3.733
B2	5.906	5.422	5.197	4.966	4.570	4.047	3.840	3.465
B2.5	5.892	5.384	5.144	4.897	4.479	3.927	3.716	3.332
B3	5.879	5.348	5.091	4.829	4.387	3.806	3.590	3.199
B4	5.854	5.280	4.991	4.696	4.206	3.562	3.335	2.936
B5	5.833	5.220	4.898	4.569	4.030	3.319	3.078	2.681
B6	5.815	5.167	4.814	4.451	3.860	3.080	2.824	2.436
B7	5.802	5.123	4.739	4.342	3.700	2.849	2.576	2.203
B8	5.791	5.087	4.674	4.244	3.550	2.628	2.337	1.985
B9	5.782	5.060	4.619	4.158	3.414	2.420	2.110	1.783
A0	5.773	5.039	4.574	4.083	3.291	2.228	1.898	1.596
A1	5.768	5.029	4.549	4.039	3.216	2.109	1.766	1.481
A2	5.761	5.021	4.528	4.000	3.149	1.999	1.641	1.373
A4	5.745	5.012	4.496	3.937	3.033	1.803	1.416	1.177
A5	5.735	5.009	4.484	3.911	2.984	1.718	1.316	1.089
A6	5.724	5.008	4.475	3.890	2.942	1.642	1.224	1.005
A8	5.695	5.007	4.464	3.859	2.876	1.517	1.067	0.853
F0	5.658	5.007	4.459	3.841	2.832	1.428	0.943	0.717
F2	5.613	5.007	4.458	3.833	2.809	1.373	0.852	0.590
F4	5.564	5.006	4.461	3.834	2.805	1.350	0.791	0.469
F5	5.539	5.005	4.464	3.837	2.809	1.350	0.771	0.409
F6	5.513	5.004	4.466	3.841	2.817	1.358	0.758	0.349
F8	5.465	5.003	4.473	3.854	2.842	1.395	0.750	0.226
G0	5.424	5.005	4.482	3.872	2.880	1.456	0.766	0.098
G2	5.408	5.007	4.488	3.882	2.903	1.496	0.782	0.031
G4	5.395	5.011	4.494	3.894	2.928	1.541	0.802	-0.038
G8	5.385	5.025	4.510	3.919	2.985	1.645	0.857	-0.183
K0	5.388	5.035	4.520	3.933	3.016	1.703	0.890	-0.258
K1	5.397	5.048	4.531	3.948	3.048	1.766	0.926	-0.336
K3	5.435	5.083	4.558	3.980	3.116	1.899	1.007	-0.497
K4	5.465	5.106	4.574	3.998	3.151	1.969	1.050	-0.580
K5	5.501	5.133	4.593	4.017	3.187	2.041	1.095	-0.665
K7	5.595	5.196	4.635	4.057	3.260	2.188	1.188	-0.839
K9	5.715	5.275	4.686	4.100	3.334	2.336	1.281	-1.019
M0	5.784	5.320	4.715	4.124	3.370	2.409	1.325	-1.111
M1	5.933	5.421	4.780	4.176	3.444	2.550	1.407	-1.299
M2	6.092	5.536	4.858	4.236	3.520	2.684	1.475	-1.492
M3	6.251	5.666	4.952	4.309	3.601	2.808	1.523	-1.692
M4	6.401	5.812	5.068	4.405	3.697	2.925	1.548	-1.901
M5	6.531	5.981	5.218	4.537	3.820	3.040	1.547	-2.122
M6	6.635	6.185	5.422	4.727	3.993	3.166	1.520	-2.356
M7	6.710	6.445	5.707	5.009	4.247	3.321	1.473	-2.604

Table B.7: *Luminosity calibration of the MK spectral types for different luminosity classes. The logarithm of L is given in $L_\odot$. Source: de Jager & Nieuwenhuijzen (1987, A&A 177, 217). In the article a fit-formula is given that provides the luminosity as a function of a numerical spectral type and luminosity class. This formula (based on a 20th order Chebychev polynomial) is not used to generate the tabels provided here and in the original paper (that are based on a 40th order Chebychev polynomial).*

$T_{\rm eff}$	$\log g$	U	B	V	R	I	J	H	K
3500	3.50	0.8051	0.8536	0.8265	0.7515	0.6551	0.4581	0.4371	0.3618
	4.00	0.7690	0.7941	0.7736	0.7198	0.6242	0.3979	0.3409	0.2836
	4.50	0.7440	0.7573	0.7380	0.6956	0.6014	0.3664	0.2887	0.2428
	5.00	0.7252	0.7358	0.7130	0.6781	0.5853	0.3522	0.2693	0.2268
4000	3.50	0.8538	0.9011	0.8348	0.7786	0.6690	0.5307	0.4812	0.4083
	4.00	0.8194	0.8728	0.8117	0.7618	0.6549	0.5136	0.4688	0.3961
	4.50	0.7767	0.8131	0.7555	0.7132	0.6137	0.4628	0.4251	0.3545
	5.00	0.7411	0.7525	0.6947	0.6565	0.5655	0.4048	0.3554	0.2962
4500	3.50	0.9213	0.8988	0.8078	0.7272	0.6263	0.5092	0.4439	0.3793
	4.00	0.8987	0.8927	0.8040	0.7266	0.6258	0.5087	0.4443	0.3803
	4.50	0.8704	0.8837	0.7990	0.7259	0.6255	0.5078	0.4439	0.3802
	5.00	0.8359	0.8650	0.7846	0.7162	0.6178	0.4984	0.4359	0.3724
5000	3.50	0.9154	0.8588	0.7599	0.6803	0.5890	0.4676	0.3999	0.3439
	4.00	0.9045	0.8567	0.7611	0.6817	0.5898	0.4700	0.4028	0.3466
	4.50	0.8882	0.8536	0.7615	0.6830	0.5905	0.4721	0.4055	0.3491
	5.00	0.8670	0.8485	0.7605	0.6834	0.5910	0.4732	0.4064	0.3501
5500	3.50	0.8767	0.8133	0.7067	0.6297	0.5453	0.4251	0.3574	0.3108
	4.00	0.8779	0.8135	0.7097	0.6329	0.5483	0.4287	0.3622	0.3145
	4.50	0.8745	0.8122	0.7117	0.6353	0.5507	0.4315	0.3659	0.3171
	5.00	0.8658	0.8097	0.7132	0.6375	0.5531	0.4340	0.3687	0.3190
6000	3.50	0.8107	0.7687	0.6581	0.5819	0.5003	0.3851	0.3172	0.2765
	4.00	0.8192	0.7682	0.6608	0.5857	0.5051	0.3899	0.3236	0.2818
	4.50	0.8250	0.7673	0.6630	0.5887	0.5090	0.3937	0.3292	0.2860
	5.00	0.8267	0.7653	0.6640	0.5904	0.5117	0.3959	0.3327	0.2883
6500	3.50	0.7410	0.7330	0.6236	0.5459	0.4642	0.3529	0.2850	0.2502
	4.00	0.7497	0.7292	0.6233	0.5478	0.4682	0.3572	0.2913	0.2551
	4.50	0.7586	0.7263	0.6233	0.5494	0.4719	0.3609	0.2973	0.2593
	5.00	0.7661	0.7226	0.6222	0.5496	0.4742	0.3631	0.3021	0.2622
7000	3.50	0.6849	0.7070	0.6016	0.5207	0.4364	0.3275	0.2610	0.2319
	4.00	0.6894	0.7003	0.5997	0.5215	0.4403	0.3315	0.2661	0.2358
	4.50	0.6959	0.6942	0.5974	0.5216	0.4435	0.3346	0.2715	0.2393
	5.00	0.7051	0.6894	0.5957	0.5221	0.4472	0.3383	0.2777	0.2430
7500	3.50	0.6487	0.7111	0.6067	0.5157	0.4214	0.3147	0.2490	0.2204
	4.00	0.6397	0.6761	0.5793	0.4969	0.4131	0.3073	0.2448	0.2192
	4.50	0.6471	0.6690	0.5790	0.5003	0.4204	0.3132	0.2517	0.2242
	5.00	0.6541	0.6621	0.5765	0.5004	0.4244	0.3168	0.2573	0.2275

Table B.8: *Limb darkening coefficients for Kurucz model atmospheres for solar-type stars. The limb darkening is described by a linear law:* $I_m(\mu)/I_m(1) = 1 - u(1-\mu)$*, where* m *is the label for the photometric passband. The models have a solar chemical composition and are for a micro-turbulent velocity of 2* $\mathrm{km\,sec^{-1}}$ *(Claret 2000, A&A 363, 1081).*

Table B.9: *Kurucz-model atmospheres for solar composition stars*

$T_{\rm eff}$ log g	log τ	log z	T	log $p_{\rm G}$	log $n_{\rm e}$	log N_N	log ρ	log $p_{\rm R}$	$F_{\rm conv}/F$
5 500	-3.0	6.79	4 282	3.23	11.35	15.47	-8.19	0.09	0.00
4	-2.0	7.65	4 487	3.84	11.91	16.05	-7.61	0.10	0.00
	-1.0	7.92	4 846	4.41	12.49	16.59	-7.07	0.17	0.00
	0.0	8.08	6 130	4.92	13.50	16.99	-6.66	0.54	0.01
	1.0	8.14	8 176	5.10	14.94	17.04	-6.62	1.05	0.85
5 500	-3.0	10.65	4 104	1.28	9.53	13.52	-10.13	0.09	0.00
1	-2.0	10.98	4 444	2.09	10.35	14.30	-9.36	0.10	0.00
	-1.0	11.14	4 846	2.73	11.08	14.91	-8.75	0.17	0.00
	0.0	11.22	6 145	3.13	12.55	15.20	-8.46	0.56	0.00
	1.0	11.24	8 431	3.18	14.06	15.07	-8.58	1.10	0.91
6 000	-3.0	7.60	4 667	3.29	11.48	15.49	-8.17	0.24	0.00
4	-2.0	7.90	4 891	3.87	12.04	16.04	-7.61	0.25	0.00
	-1.0	8.08	5 293	4.42	12.62	16.55	-7.10	0.32	0.00
	0.0	8.18	6 789	4.82	13.94	16.85	-6.81	0.70	0.05
	1.0	8.22	8 709	4.95	15.12	16.86	-6.79	1.16	0.88
6 000	-3.0	10.75	4 489	1.26	9.72	13.47	-10.19	0.24	0.00
1	-2.0	11.03	4 869	2.02	10.62	14.19	-9.47	0.25	0.00
	-1.0	11.17	5 318	2.59	11.44	14.72	-8.94	0.33	0.00
	0.0	11.24	6 861	2.89	13.01	14.90	-8.75	0.75	0.00
	1.0	11.25	8 981	2.92	14.11	14.73	-8.93	1.21	0.91
7 000	-3.0	7.63	5 458	3.10	11.87	15.22	-8.44	0.51	0.00
4	-2.0	7.95	5 726	3.67	12.45	15.77	-7.89	0.52	0.00
	-1.0	8.12	6 190	4.17	13.13	16.23	-7.42	0.60	0.00
	0.0	8.20	8 217	4.45	14.63	16.39	-7.26	1.02	0.20
	1.0	8.24	9 911	4.55	15.37	16.37	-7.28	1.38	0.92
10 000	-3.0	8.34	7 586	1.71	12.84	13.63	-10.03	1.13	0.00
4	-2.0	8.48	8 030	2.36	13.42	14.36	-9.40	1.15	0.00
	-1.0	8.58	8 982	2.86	14.08	14.67	-8.99	1.28	0.00
	0.0	8.65	11 655	3.17	14.62	14.71	-8.95	1.68	0.00
	1.0	8.83	16 287	3.75	15.08	15.12	-8.54	2.25	0.00
20 000	-3.0	8.70	13 060	1.38	12.81	12.84	-10.82	2.34	0.00
4	-2.0	8.90	14 067	2.09	13.49	13.52	-10.14	2.35	0.00
	-1.0	9.02	15 560	2.71	14.07	14.08	-9.57	2.40	0.00
	0.0	9.15	19 521	3.33	14.60	14.60	-9.05	2.63	0.00
	1.0	9.28	27 451	4.03	15.15	15.15	-8.50	3.15	0.00
40 000	-3.0	9.48	28 059	1.19	12.31	12.29	-11.37	3.54	0.00
4	-2.0	9.66	31 336	2.16	13.24	13.21	-10.45	3.55	0.00
	-1.0	9.77	34 855	2.93	13.96	13.93	-9.72	3.62	0.00
	0.0	9.87	40 920	3.55	14.52	14.48	-9.18	3.85	0.00
	1.0	9.97	53 682	4.21	15.06	15.02	-8.64	4.32	0.00

τ = continuum optical depth at λ5000 Å; z = geometrical depth in cm; T = temperature in K; $p_{\rm G}$ = gas pressure; $n_{\rm e}$ = electron density in cm^{-3}; N_N = nucleon density in cm^{-3}; ρ = density in gr cm^{-3}; $p_{\rm R}$ = radiation pressure; $F_{\rm conv}/F$ = fraction of the flux that is transported by convection. All units in cgs. From: Cox (ed.): Allen's Astrophysical Quantities, 4th edition, section 15.4.1., p. 393.

λ Å	W_λ mÅ	Elem. & Ion	Fr. id	λ Å	W_λ mÅ	Elem. & Ion	Fr. id	λ Å	W_λ mÅ	Elem. & Ion	Fr. id
2795.4		Mg II		3709.256	573	Fe I[1]		4077.724	428	Sr II[1]	
2802.3		Mg II		3719.947	1664	Fe I		4101.748	3133	Hδ	h
2851.6		Mg		3734.874	3027	Fe I	M	4132.067	404	Fe I[1]	
2881.1		Si		3737.141	1071	Fe I		4143.878	466	Fe I	
3067.262	663	Fe I[1]		3745.574	1202	Fe I[1]		4167.277	200	Mg I	
3134.116	414	Ni I[1]		3748.271	497	Fe I		4202.040	326	Fe I	
3242.007	270	Ti II		3749.495	1907	Fe I		4226.740	1476	Ca I	g
3247.569	246	Cu I		3758.245	1647	Fe I		4235.949	385	Fe I[1]	
3336.689	416	Mg I		3759.299	334	Ti II		4250.130	342	Fe I[1]	
3414.779	816	Ni I		3763.803	829	Fe I		4250.797	400	Fe I[1]	
3433.579	492	Ni I[1]		3767.204	820	Fe I		4254.346	393	Cr I[1]	
3440.626	1243	Fe I		3787.891	512	Fe I		4260.486	595	Fe I	
3441.019	634	Fe I		3795.012	547	Fe I[1]		4271.774	756	Fe I	
3443.884	655	Fe I		3806.718	209	Fe I[1]		4325.775	793	Fe I[1]	
3446.271	470	Ni I		3815.851	1272	Fe I		4340.475	2855	Hγ	f
3458.467	656	Ni I		3820.436	1712	Fe I	L	4383.557	1008	Fe I	e
3461.667	758	Ni I		3825.891	1519	Fe I		4404.761	898	Fe I	
3475.457	622	Fe I		3827.832	897	Fe I		4415.135	417	Fe I[1]	
3476.712	465	Fe I[1]		3829.365	874	Mg I		4528.627	275	Fe I[1]	
3490.594	830	Fe I		3832.310	1685	Mg I		4554.036	159	Ba II	
3492.975	826	Ni I		3834.233	624	Fe I		4703.003	326	Mg I	
3497.843	726	Fe I		3838.302	1920	Mg I		4861.342	3680	Hβ	F
3510.327	489	Ni I		3840.447	567	Fe I		4891.502	312	Fe I	
3515.066	718	Ni I		3841.058	517	Fe I[1]		4920.514	471	Fe I[1]	
3521.270	381	Fe I		3849.977	608	Fe I		4957.613	696	Fe I[1]	c
3524.536	1271	Ni I		3856.381	648	Fe I		5167.327	935	Mg I[1]	b_4
3554.937	404	Fe I		3859.922	1554	Fe I		5172.698	1259	Mg I	b_2
3558.532	485	Fe I[1]		3878.027	555	Fe I		5183.619	1584	Mg I	b_1
3565.396	990	Fe I		3886.294	920	Fe I		5250.216	62	Fe I[1]	
3566.383	458	Ni I		3899.719	436	Fe I		5269.550	478	Fe I[1]	E
3570.134	1380	Fe I		3902.956	530	Fe I[1]		5328.051	375	Fe I	
3578.693	488	Cr I		3905.532	816	Si I		5528.418	293	Mg I	
3581.209	2144	Fe I	N	3920.269	341	Fe I		5889.973	752	Na I[1]	D_2
3586.990	532	Fe I		3922.923	414	Fe I[1]		5895.940	564	Na I	D_1
3593.495	436	Cr I		3927.933	187	Fe I		6102.727	135	Ca I	
3608.869	1046	Fe I		3930.308	108	Fe I		6122.226	222	Ca I	
3618.777	1410	Fe I		3933.682	20253	Ca II[1]	K	6162.180	222	Ca I	
3619.400	568	Ni I		3944.016	488	Al I		6302.499	83	Fe I[1]	
3631.475	1364	Fe I		3961.535	621	Al I		6562.808	4020	Hα	C
3647.851	970	Fe I		3968.492	15467	Ca II[1]	H	8498.062	1470	Ca II	
3679.923	448	Fe I		4045.825	1174	Fe I		8542.144	3670	Ca II	
3685.196	275	Ti II		4063.605	787	Fe I[1]		8662.170	2600	Ca II	
3705.577	562	Fe I		4071.749	723	Fe I		10830		He I	

Table B.10: *The most important Fraunhofer lines from the Sun after Moore, Minnaert, and Houtgast (1966).* [1] *Blended line. Adapted from: Lang (1980), 2*nd *edition, page 175. The column labeled Fr provides the letter designation given originally by Fraunhofer to the most prominent absorption features. From: Gray, 3*rd *edition, page 521.*