AE4-805: Astronomy

docent: F.P. Israel

12 September 2003

Delft University of Technology

AE4-805: Astronomy/docent: F.P. Israel/12 September 2003

This document was typeset and translated with permission of the author by R.A. Deurloo at the research group

Astrodynamics & Satellite Systems
Faculty of Aerospace Engineering
Delft University of Technology
Kluyverweg 1, 2629 HS Delft
The Netherlands

The document was typeset with LaTeX 2_ε. The layout is an adaptation (based on ideas by Marvin Stiefelhagen) of the 'thesis.cls' layout originally designed by Remco Scharroo

AE4-805: Astronomy

by

docent: F.P. Israel

Contents

Chapter 1
Introduction

1.1 The Science of Astronomy

Astronomy is the oldest science in the world. It is also the only science to which the ancient Greeks assigned a Muse: Urania. For millennia, astronomy was tainted by religious and metaphysical prejudice; it even found its justification in those very aspects. Only in the 17^{th} century astrology and astronomy progressively departed ways. In the centuries following, astronomy was largely limited to the determiniation of position and motion of the planets and stars. Its nature was observational and mathematical. In the 19^{th} century, the branch of astrophysics developed, focussing on the physics of processes in space. In the 20^{th} century astrochemistry developed, and perhaps the 21^{th} century will bring astrobiology.

Astronomy is an experimental science. However, the experiment is performed by nature, not by us. *All* astronomical knowledge is based on the accurate *observation* of electromagnetic radiation received from space (sometimes after traveling for millions or billions of years) and on the *interpretation* of these observations with the aid of other sciences such as experimental and theoretical physics, mathematics and chemistry. Visible light is only a small part of the entire spectrum of electromagnetic radiation emitted by objects in space. Ever since the work by Christiaan Huygens and by Isaac Newton in the 17^{th} century, we know that electromagnetic radiation may be considered as a wave-phenomenon, but also as a stream of particles called photons. Confusingly, these *mutually exclusive* descriptions are fully *complementary*: any description of electromagnetic radiation employing only one of the two concepts is incomplete. Electromagnetic radiation transports energy: the energy E, carried by a photon with frequency ν is $E = h\nu$.

Many (but not all!) astronomical processes take place on characteristic time scales that vastly exceed the human lifespan, and may even be longer than the thousands of years of written human history. As a result, astronomical insights are mostly gained by considering statistically significant, large numbers of objects with different properties and in different evolutionary stages.

Practical astronomy deals with the determination of *intensities* as a function of (i). position in the sky, (ii). wavelength or frequency, (iii). polarization and (iv). time. It attempts to find patterns and order in the observational data gathered. The experimental character of astronomy means that is often hard to predict in advance what information will be of real value and what will turn out to be trivial or superfluous. Using physics and mathematical techniques, astronomical *theory* tries to fit observed facts and correlations into a conceptual framework meant to advance understanding of the whole. In reality, however, the partition between practical and theoretical astronomy is not so clear-cut. Most astronomical research consists of both practical and theoretical components. It is only the balance between the two that differs per investigation.
In astronomy, a broad array of techniques is used:

- **Astrometry**: accurate determination of the position of an object in the sky. The atmosphere is a limiting factor (seeing) as are imperfections in the mechanical properties of telescopes. The ESA satellite Hipparcos, has accurately documented position and motion of more than 100000 stars, providing results ten to a hundred times better than obtained from the ground.
- **Photometry**: accurate measurement of the brightness or intensity of objects in the sky in relatively wide pass-bands (defined by filters) centered at different wavelengths ('colours') as well as brigthness variation as a function of time. Imaging might be considered a very general form of photometry.
- **Spectroscopy**: determination of the intensity distribution of objects as a function of wavelength (frequency) with relatively high spectral resolutions. A hybrid form of spectroscopy and photometry is spectrophotometry, with relatively low spectral resolutions of the order of a few percent of the wavelength.
- **Polarimetry**: measurement of the type, degree and angle of polarization of the radiation from objects in the sky.

1.2 The Electromagnetic Spectrum

Depending on the experiment carried out, radiation must be viewed either as a stream of particles (photons) with rest-mass $m = h\nu/c^2$, or as an electromagnetic wave propagating through space with perpendicular electric and magnetic oscillations. In a vacuum, radiation propagates at an invariable speed ($c = 2.997925 \times 10^8$ m s^{-1}) usually called the speed of light. In other media, the actual propagation speed is somewhat lower.

The totality of all possible frequencies ν or wavelengths λ (which are equivalent: $c = \lambda \times \nu$) is called the *electromagnetic spectrum*. Traditionally, one distinguishes different regions of the electromagnetic spectrum as follows:

- *Radio region*: wavelengths of 1 mm and larger; subdivided into a microwave region (between 1 mm and 1 dm; EHF, SHF); an (ultra) short-wave region (between 1 dm and 10 m: UHF, VHF); a medium-wave region (between 10 m and 1000 m: HF, MF), and a long-wave region (larger than 1 km; LF, VLF, ULF).
- *Infrared region*: wavelengths between $1\ \mu\mathrm{m} = 0.001$ mm and 1 mm; subdivided into a submillimetre region (between $200\ \mu$m and 1 mm), a far-infrared region (FIR, between 50 and $200\ \mu$m), a mid-infrared region (MIR, between 10 and $50\ \mu$m), and a near-infrared region (NIR, between 1 and $10\ \mu$m).
- *Visual region*, also known as optical region: between 3000 Angström $= 0.3\ \mu$m and $1\ \mu$m; subdivided into (infra)red, yellow and blue.
- *Ultraviolet region*: between 3000 Angström and approx. 300 Angström; subdivided into near-ultraviolet (NUV), ultraviolet (UV) and far-ultraviolet (FUV).
- *X-ray region*: also called Röntgen region, between 300 Angström and 0.3 Angström; subdivided into 'soft' (long-wave) and 'hard' (short-wave) X-rays.
- *Gamma-ray region*: smaller than 0.3 Angström.

The colour of a star provides a crude estimates of its temperature. The colour of the brightest stars can even be seen with the unaided eye). In Table 1.1 we show stellar temperatures, wavelengths of peak light emission, and the blue-to-red light intensity ratio – commonly used to define stellar colours.

As we will see later, the temperature of objects determines in what part of the electromagnetic spectrum they radiate most, and consequently in what part of the spectrum they are best observed. For example, colour and temperature of very hot stars ($T > 20000$ K) are hard to measure in visual light. Most of the radiation is emitted in the UV. They also have their most interesting (absorption) lines in the UV. In contrast, very cool objects

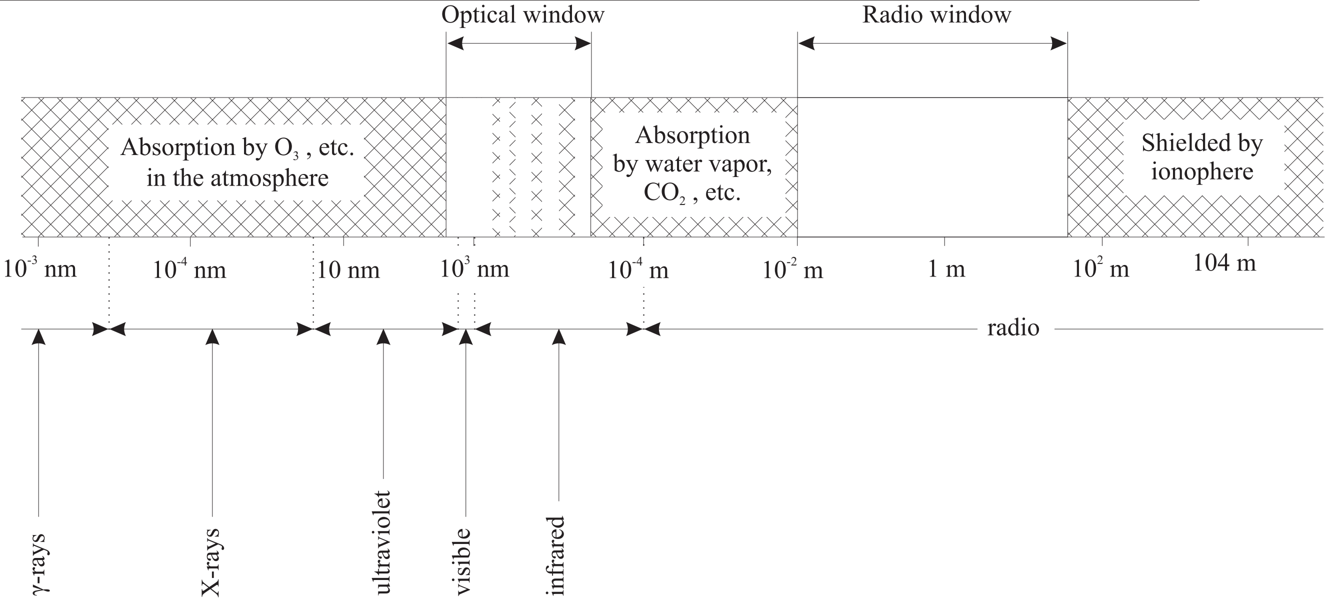

Figure 1.1 The electromagnetic spectrum and the 'windows' of the Earth's atmosphere. From: Gathier Sterrekunde.

Temperature (K)	λ_{max} (Å)	I_{blue}/I_{red}	'Colour'
30000	965	3.6	far UV (blue)
20000	1450	3.2	near UV (blue)
10000	2900	2.2	blue
6000	4835	1.1	white
4000	7250	0.54	red
3000	9665	0.26	IR (deep red)
2000	14500	0.068	IR (near infra-red)

Table 1.1 Temperatures with corresponding peak intensity wavelength, blue and red light ratio, and 'colour'.

($T << 2000$ K) radiate little, if at all in the visual, but are obvious at infrared wavelengths. Such cool objects may have very informative molecular lines in the infrared and sub-millimeter parts of the spectrum. Unfortunately, the terrestrial atmosphere transmits poorly in the IR and UV, so that space-based satellite observatories offer great advantages, if not the only way of performing the desired measurements.

1.3 Angular Resolution

The theoretical resolution of any optical system (eye, telescope, radio telescope) is determined by *diffraction*. The diffraction pattern of a pointlike source is often referred to as the *Point Spread Function* (PSF). Primarily, it is a function of the wavelength λ of the observed radiation and the aperture D (or maximum baseline D in case of an interferometer) of the imaging system. Numerically:

$$\sin\phi = 1.22\lambda/D\,, \tag{1.1}$$

where ϕ (in degrees) is the *resolvable angle* or (angular) *resolution*. The resolution ϕ corresponds to the minimum angular distance under which two point sources can still be separated.

Example *Under normal conditions, the pupil of an eye is approx.* 2 mm *in diameter. The retina detects light at a wavelength of* 5000 Å*. The human eye has a resolution* ϕ *of about* $1'$

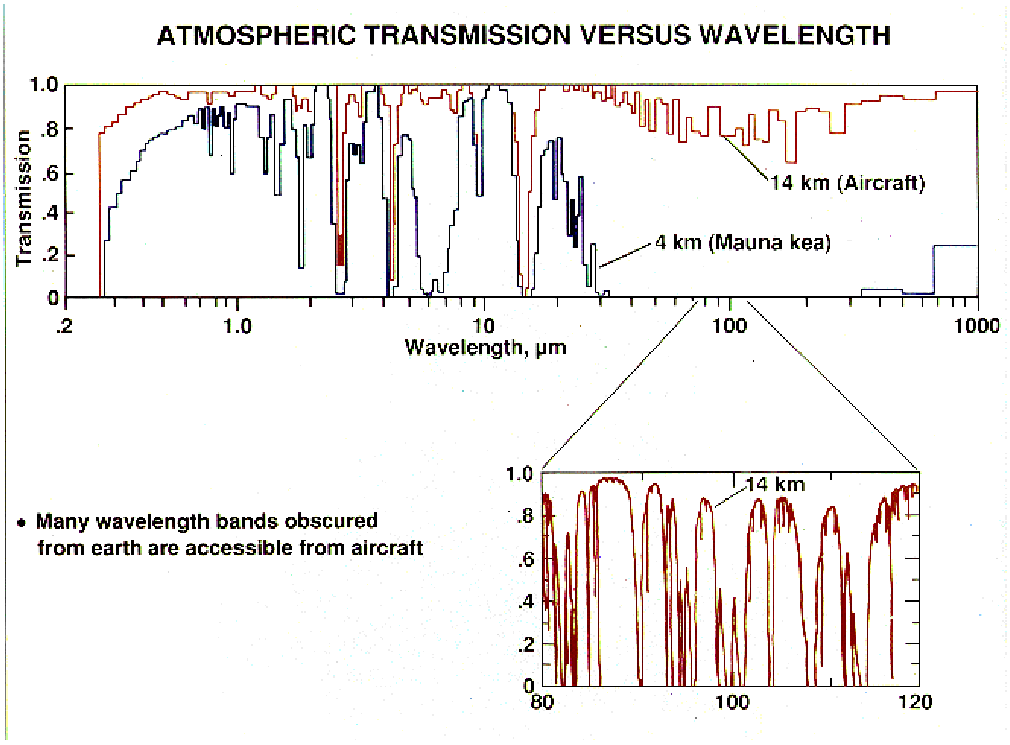

Figure 1.2 Atmospheric transmission at altitudes of 4 and 14 km. Taken from *SOFIA: The Future of Airborne Astronomy* by E.F. Erickson and J.A. Davidson, in *Proc. of the Airborne Astron. Symp. on the Galactic Ecosystem*, Eds. M.R. Haas, J.A. Davidson & E.F. Erickson (San Francisco: ASP). Also at: http://sofia.arc.nasa.gov/Science/publications/erickson

($\sin\phi = 3.0 \times 10^{-4}$). At a wavelength of 21.1 cm, the resolution of the huge 25 m Dwingeloo radio telescope is only $35'$. However, at the same wavelength the 3.2 km Westerbork Synthesis Radio Telescope (WSRT) has a resolution $\phi = 16''$. Thus, the vision of the naked eye is about equal tot that of an 800 m radio telescope!

Modern techniques (*adaptive optics*) allow very fast adaptation of telescope mirror shapes in order to compensate for more slowly changing atmospheric disturbances. With important limitations, groundbased telescopes thus equipped may likewise approach their theoretical resolving power (ESO-VLT in Chili; California-Keck Telescope on Hawaii).

1.4 Telescopes

An imaging system (telescope) brings together elements of a wavefront by ensuring that all elements have precisely the same travel time. If mirrors are used, this is the case if the geometrical pathways through the telescope are all equally long (i.e. all rays reflected by a parabolic mirror go through the same point). If lenses are used, longer geometrical pathways are compensated by lower propagation speeds inside the lens. Nowadays, parabolic mirrors are the primary radiation-collecting devices over wavelengths ranging from 300 Angström = 30 nm (UV) to about 1 m (radio). Towards shorter wavelengths, mirror-surfaces must be of ever-increasing accuracy. A radio telescope may be equipped with a coarse mesh in a parabolic shape no more accurate than a few millimeters, whereas an optical/UV telescope requires very precisely polished surfaces with Angström accuracies.
Parabolic mirrors may have greatly varying focal lengths, usually expressed as *focal ratios f/x*, where x is the numerical ratio between focal length and aperture diameter.

- *Prime Focus.* Directly in front of the primary mirror. Detectors placed here block part of the incoming radiation and deform the image. It is not used much in optical

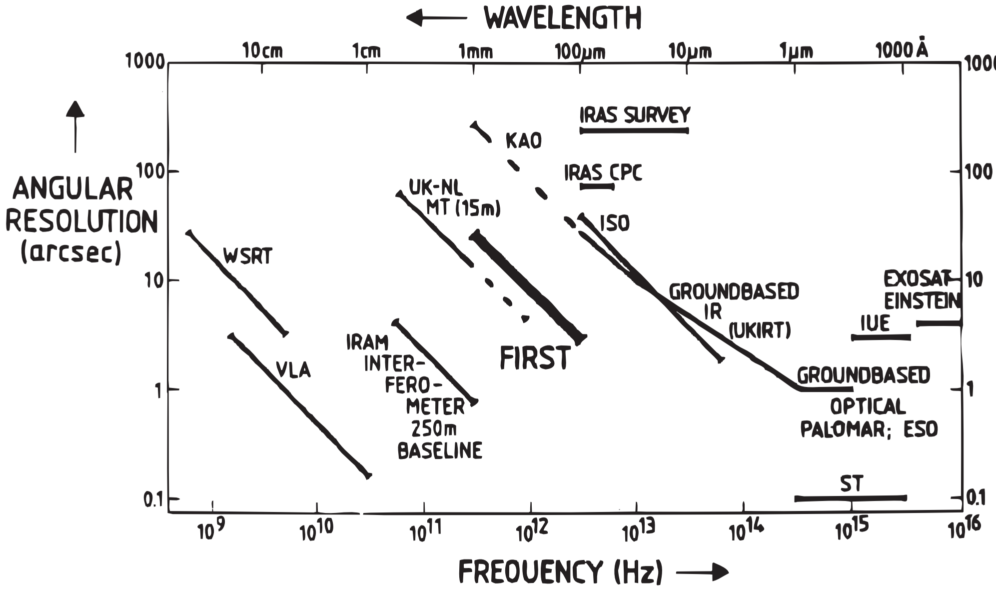

Figure 1.3 The angular resolutions of some space-based and ground-based telescopes as a function of wavelength (scale at the top) and frequency (scale at the bottom). From ESA FIRST proposal

astronomy but it is still the preferred radio telescope focus. Typical focal ratio f/1.5 - f/3.

- *Newton Focus.* Is really a modified prime focus obtained by placing a flat mirror tilted by 45° between primary mirror and prime focus, throwing the actual focus to a position outside the the path of the incoming radiation. Popular with amateur astronomers, because only one mirror with a curved surface is required.
- *Cassegrain Focus.* A concave secondary mirror, placed between primary mirror and prime focus, projects a lengthened focus back to a point behind the primary mirror, which thus must have a central hole to let the light through. Allows bulky detector systems to be placed on the telescope. Most frequently used focal arrangement in professional astronomy, also in spacecraft. Typical focal ratio f/8 - f/15.
- *Nasmyth Focus*. Is really a modified Cassegrain focus. As in that case, radiation is bounced back by a concave secondary mirror. Subsequently, as in the Newton case, it is reflected out of the main light pathway along the telescope elevation axis by a flat tertiary mirror tilted by 45°. The mirror can be flipped to left or right, allowing even more bulky detector systems to be placed at two different, fixed locations next to the telescope. Becoming very popular, especially for radio telescopes.
- *Coudé Focus*. An extremely long Cassegrain/Nasmyth focus located well outside the telescope (or even telescope building) is obtained by applying a sequence of additional mirrors. Used for spectroscopy at very high spectral resolution. Unfit for two-dimensional imaging. Typical focal ratio f/35 - f/100.

X-ray telescopes need a different design, because wavelengths are so short (and energies so high) that radiation at this wavelengths will penetrate parabolic mirrors of any material and be absorbed. However, total reflection of rays still occurs for grazing (small angle) incidence. Unlike optical imaging systems, which can be 'folded' by using back-reflecting mirrors, X-ray mirrors allow only forward reflections. X-ray (space) telescopes consist of an arrangement of parabolic and hyperbolic ringlike mirror segments, and are quite tall.

Gamma-ray telescopes cannot even use grazing-incidence reflection. Rather, they make use of the Compton effect. A high-energy gamma photon colliding with a loosely bound electron transfers a fraction of its energy to that electron which then moves off under an angle ϕ with the original direction of travel of the gamma photon. The direction in which the electron moves can be detected, providing us with a circle on the sky of radius

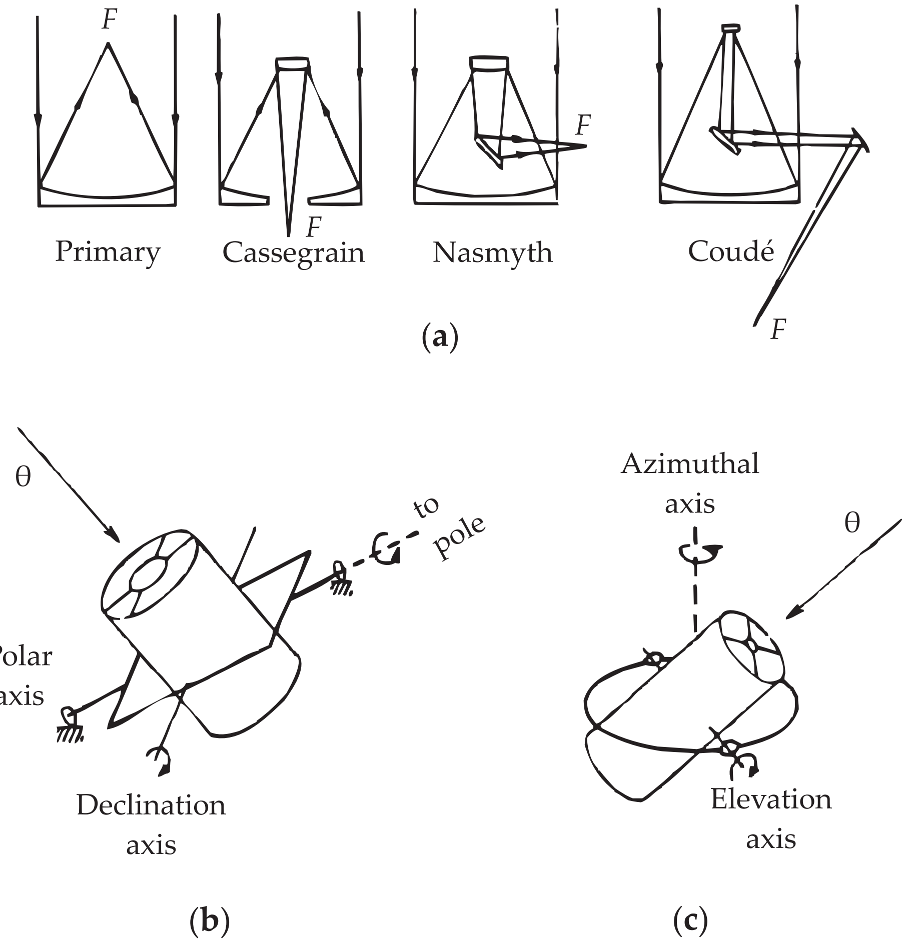

Figure 1.4 Top: Frequently used focus configurations of astronomical telescopes. Bottom: Axis arrangements of groundbased telescopes. From: P. Lena, F. lebrun & F. Mignard: *Observational Astrophysics*, Springer, Berlin 1998

ϕ on which the source of the gamma photon must lie. Multiple observations of a gamma source may provide additional circles with a single point of overlap.

1.5 Detectors

After imaging, radiation must be detected. Of old, the retina of the human eye was used as a detector. Its major disadvantage, lack of storage capacity, was solved by the introduction of the photographic emulsion. Major drawbacks of the emulsion are its mostly non-linear response (in underexposed and overexposed regimes respectively), and its low quantum (photon) efficiency. After being a major tool of astronomy for over a century, emulsions are no longer in general use. Their place has been taken by:

- **Bolometers**, semiconductors cooled by liquid nitrogen or liquid helium. Incident radiation causes a small increase in bolometer temperature, leading to a reduced resistance to electric current. The decrease in resistance can be measured accurately. Works over a very large wavelength range. Much used in the infrared part of the electromagnetic spectrum.
- **Photon Detectors**, semiconductors in which electrons are excited by incident photons causing either (a). a voltage increase (photovoltaic detector) or (b). a resistance decrease (photoconductor) which can be detected. However, there is an energy threshold for electron excitation, so that photon detectors only work *below* a critical wavelength.

- **Charge-Coupled Device (CCD)**, ever-larger chips containing semiconductors (pixels) in which electrons are excited by incident photons; excited electrons are then counted. CCD's have very high quantum efficiencies (close to 100%), linear responses and wide dynamic range. They do suffer from read-out noise and uneven pixel response. They must be cooled and are sensitive to cosmic ray hits (especially in space). They can be used to produce images in the range of 200 (UV) to 1100 nm (NIR), and at X-ray wavelengths. At longer wavelengths, photons lack the energy to excite the electrons.

1.6 Importance of Space Observatories

The Earth's atmosphere and ionosphere are transparent to a very small fraction of the electromagnetic spectrum. Only *visual radiation, VHF/UHF/SHF radio radiation, NIR and submillimetre* radiation can reach the Earth's surface through a few atmospheric 'windows'. Radiation at any other wavelength can only be detected by spacecraft outside the Earth's atmosphere. This is particularly important for radiation in *the gamma, X and ultraviolet regions, most of the infrared, and the long-wavelength part of the radio region* (> 35 m).

- **Radiation in near-infrared, mid-infrared, far-infrared and submillimetre regions** is effectively absorbed by molecules in the atmosphere (in particular H_2, O_2, O_3 and CO_2). This absorption is not continuous, but occurs in often overlapping molecular absorption bands. In the wavelength range from 2 μm (150 GHz) to 1 cm (30 GHz), only a few small spectral regions are (mostly) free of this absorption. Such 'free' regions of relatively good transparency are called ' atmospheric windows'. In these windows, groundbased observations are posible. Outside these windows space-based satellites are required (IRAS, ISO, SIRTF, Herschel).
- **Radio radiation** at wavelengths of > 15 m (< 20 MHz) or longer is completely blocked reflected by the ionosphere. For these wavelengths one has to use satellites (Radio Astronomy Explorer, RAE). The *interplanetary medium* sets a further, effective limit at wavelengths > 100 m (< 3 MHz).
- **UV radiation** is absorbed, mainly by ozone (O_3), in the outer layers of the atmosphere – fortunate for humans who now only get a modest sunburn. Thus, only satellites can perform astronomical observations at wavelengths < 3000 angström (ANS, Copernicus, Far-UV and Extreme-UV Explorers).
- **X-rays and gamma-rays** are completely absorbed by the dense layers of the atmosphere. Sources of high-energy of radiation need to be observed at wavelengths in the extreme UV, X-ray or gamma-ray regions. This radiation is essential to our understanding of processes which take place in exotic objects such as supernovae, supernova remnants, accretion disks near black holes, galaxy nuclei, quasars, superhot gas between galaxies etc. (satellite observatories: X-rays: Uhuru, ExoSat, RoSat, Chandra, XMM-Newton; Gamma-rays: Cos-B, GRO, Integral).

Additional reasons for choosing a satellite to perform astronomical observations:

- **Infrared** observations of cosmic objects are hampered by the fact that the Earth and everything on it (including telescopes and detectors) emits radiation at these same wavelengths: for $T = 300$ K maximum radiation is at 10 μm. For this reason, infrared detectors are cooled to a few K (liquid helium cooling). On Earth, the telescope itself cannot be cooled efficiently because of convective and conductive thermal losses. However, in space this is possible because of the absence of such losses in a vacuum. Current technologies allow the construction of dewars ('thermos flasks') that may encompass telescopes up to 1-m mirror size (IRAS, ISO: 60 cm; SIRTF: 85 cm).
- **(Visual) radiation** travelling through a turbulent atmosphere is subject to 'seeing'. It loses wavefront-coherence resulting in distorted and scintillating images much less 'sharp' than to be expected from the theoretical resolution of the telescope (see Sect. 1.3). Groundbased telescopes will not attain visual image sharpnesses better than about $0.5''$, regardless of their size. A telescope in space (first proposed by

H. Oberth in 1924) is outside the atmosphere and does not have this limitation. It should, in principle, operate at its theoretical resolution (Hubble Space Telescope).

- **Radio waves** have sufficiently long wavelengths that phase-coherent detection is possible, allowing interferometric observations to greatly increase resolving power. Attainable resolutions depend on the (maximum) baseline expressed in wavelengths. On Earth, maximum baselines are limited by the Earth's diameter. By introducing space-based antennae, much greater baselines, hence much higher resolutions are possible (VSOP)

Chapter 2

General Properties of Radiation

2.1 Flux, Radiation Intensity and Energy Density

(Note: the contents of this Chapter are treated in more detail in: *Radiative Processes in Astrophysics*, by G.B. Rybicki and A.P. Lightman (Ed. John and Sons, New York)
The energy leaving each second a spherical radiating surface with radius r_1 (e.g. a star) is:

$$E_1 = F_1\, 4\,\pi\, r_1^2\,,$$

where F_1, the energy flow per second per m^2, is called *flux*. The flux through a concentric, larger spherical surface (radius $r_2 > r_1$) is:

$$E_2 = F_2\, 4\,\pi\, r_2^2\,.$$

According to the law of energy conservation (1^{st} law of thermodynamics) $E_1 = E_2$. Thus:

$$F_1/F_2 = (r_2/r_1)^2\ ,$$

or more general:

$$F = \mathrm{cst}\ r^{-2}\,. \tag{2.1}$$

The energy flow per second per m^2 or flux decreases as the square of the distance. For example, despite an enormous energy flow at the Sun's glowing surface, this energy flow has become so diluted at the greater distance of the Earth that it produces only a comfortable temperature.

Now consider a small element dA of radiating surface. Energy leaves this surface element in a cone with solid angle $d\Omega$. The *intensity* I of the radiation is defined as the energy flow per unit time, surface and solid angle:

$$dE = I\, dt\, dA\, d\Omega\,,$$

where we express I in units $\mathrm{W\,m^{-2}\,sr^{-1}}$; a solid angle is a 2-dimensional angle with unit steradian (sr); a sphere subtends 4π sr.

For frequency (wavelength) dependent radiation we define specific intensity at frequency ν: $I_\nu = I/d\nu$. So:

$$dE_\nu = I_\nu\, dt\, dA\, d\Omega\, d\nu\ , \tag{2.2}$$

where we express I_ν in units $W\,\mathrm{m^{-2}\,sr^{-1}\,Hz^{-1}}$.

N.B.: Unlike the flux F, the intensity I, defined per solid angle, is independent of distance!

Similarly we define the *radiation energy density* u_ν per unit of volume dV:

$$dE_\nu = u_\nu(\Omega)\; dV\, d\Omega\, d\nu\,. \tag{2.3}$$

What is the connection between intensity and energy density?

Consider a cylinder of length $l = c\,dt$. c is the speed of light in the vacuum ($c = 3 \times 10^8\ \mathrm{m\,s^{-1}}$), through which the radiation propagates. The cylinder has a volume $dV = c\,dt\,dA$. The amount of radiation energy contained within the cylinder is therefore:

$$dE_\nu = u_\nu\, d\Omega\, d\nu\, c\, dt\, dA\,.$$

As the radiation within the cylinder propagates with speed c, all radiation originally inside the cylinder, has left it after a time dt has elapsed, resulting a change of the energy content of the cylinder by:

$$dE_\nu = I_\nu\, dt\, d\Omega\, d\nu\, dA\,. \qquad \text{(see Eqn. 2.2)}$$

Since the cylinder is now empty, this energy change must be equal to the original amount of radiation energy present in the cylinder (energy conservation law). Hence:

$$u_\nu(\Omega) = I_\nu / c\,. \tag{2.4}$$

Thus, energy density equals intensity divided by the speed of light.

So far we have seen how radiation propagates through a vacuum. What will be the situation in a medium that itself generates and absorbs energy? This is described by the socalled Equation of Radiative Transfer.

2.2 The Equation of Radiative Transfer

Instead of radiation leaving a surface, we now consider radiation generated within a volume (e.g. by a hot gas). The extent to which this occurs is described by the *spontaneous emission coefficient* j: the energy emitted per unit time, solid angle and volume. Definition:

$$dE_\nu = j_\nu\, dt\, d\Omega\, dV\,; \tag{2.5}$$

again: $j_\nu = j/d\nu$; units j_ν: $\mathrm{W\,sr^{-1}\,m^{-3}\,Hz^{-1}}$.

Along a pathlength ds, a beam of cross-section dA fills a volume $dV = dA\,ds$. From the previous discussion, it follows that spontaneous emission adds radiation energy to the beam with an intensity $dI = dE/dt\,dA\,d\Omega\,d\nu$, so that:

$$dI_\nu = j_\nu\, ds\,.$$

This is easy to imagine: when we look at a radiating medium (such as a gas), we see a radiation intensity that is equal to the amount emitted by a unit volume times the number of such units, i.e. the depth of the medium.

The absorption of radiation that passes through a medium can be expressed in a similar way:

$$dI_\nu = -\alpha_\nu\, I_\nu\, ds\,.$$

Here α_ν is the *absorption coefficient* (unit: $\mathrm{m^{-1}}$). The intensity decrease equals the ease with which the incoming radiation I is absorbed per unit pathlength times the number of such units, i.e. the total pathlength along which absorption occurs. Any medium will simultaneously emit and absorb, causing a total intensity change given by the sum of both processes:

$$dI_\nu/ds = -\alpha_\nu\, I_\nu + j_\nu\,. \tag{2.6}$$

This is the *equation of radiative transfer* which plays a crucial role in astrophysics. The equation underlies any attempt to calculate the propagation of radiation through any medium other than a perfect vacuum. It is of fundamental importance, whether we are considering a stellar interior, a cloud of gas and dust in interstellar space, or the atmosphere of a planet such as the Earth.

The equation of radiative transfer is a differential equation. The solution of this differential equation tells us at which intensity we will observe radiation emanating from a medium with specific properties. In considering the solution, we will distinguish between the general case and the case of a homogeneous medium, in which the coefficients j and α are independent of position s in the medium.

2.3 Solutions of the Equation of Radiative Transfer

a. *Emission only*
Rewrite:

$$dI_\nu = -\alpha_\nu\, I_\nu\, ds \,+\, j_\nu\, ds \;: \qquad \alpha_\nu = 0 \tag{2.7}$$

The solution of the radiative transfer equation is thus:

$$I_\nu\,(s) = I_\nu\,(s_o) + \int_{s_0}^{s} j_\nu\,(s')\;ds'\,, \qquad \text{(general)}$$

$$I_\nu = I_{\nu,\,o} + j_\nu\,\Delta\, s\,. \qquad \text{(homogeneous medium)}$$

The exit intensity is equal to the intensity of the incoming radiation plus the intensity added by the medium.

b. *Absorption only* Again:

$$dI_\nu = -\alpha_\nu\, I_\nu\, ds \,+\, j_\nu\, ds \;: j_\nu = 0 \tag{2.8}$$

Now the solution is:

$$I_\nu\,(s) = I_\nu(s_o)\, e^{-\int_{s_0}^{s} \alpha_\nu(s')\,ds'}\,.$$

Introduce a new variable: $d\tau_\nu = \alpha_\nu\, ds$; this *dimensionless* (verify!) variable is somewhat misleadingly called *optical depth*.
Integrate:

$$\tau_\nu\,(s) = \int_{s_0}^{s} \alpha_\nu\,(s')\;ds'\,. \tag{2.9}$$

Hence:

$$I_\nu\,(s) = I_\nu\,(s_o)\;e^{-\tau_\nu(s)}\,, \qquad \text{(general)}$$

$$I_\nu = I_{\nu,\,o}\, e^{-\tau_\nu}\,. \qquad \text{(homogeneous medium)}$$

In an absorbing medium, intensity decreases exponentially, i.e. very rapidly, with increasing pathlength.

c. *Emission and Absorption* First define the so-called *source function* $S_\nu = j_\nu/\alpha_\nu$. The source function is a property of the medium and describes the balance between the relative ease with which a medium may emit and absorb radiation.

The equation of radiative transfer is written as:

$$dI_\nu/\alpha_\nu\, ds = -I_\nu + j_\nu/\alpha_\nu\,,$$

so that:

$$dI_\nu/d\tau_\nu = -I_\nu + S_\nu \qquad \textit{or} \qquad dI_\nu = -I_\nu\, d\tau_\nu + S_\nu\, d\tau_\nu\ . \quad (2.10)$$

Integration yields the formal solution of the radiative transfer equation:

$$I_\nu\left(\tau_\nu\right) = I_\nu\left(0\right)\, e^{-\tau_\nu} + \int_0^{\tau_\nu} e^{-\left(\tau_\nu - \tau_\nu^{'}\right)}\, S_\nu\left(\tau_\nu^{'}\right)\, d\tau_\nu^{'}\ . \qquad \text{(general)}$$

For a homogeneous medium S_ν is, by definition, a constant (the value of S_ν is not dependent on the position s). As a result:

$$\begin{aligned} I_\nu\left(\tau_\nu\right) &= I_\nu\left(0\right)\, e^{-\tau_\nu} + S_\nu\left(1 - e^{-\tau_\nu}\right) \\ &= S_\nu + e^{-\tau_\nu}\left(I_\nu\left(0\right) - S_\nu\right)\ . \end{aligned} \qquad \text{(homogeneous medium)} \quad (2.11)$$

Note that for $\tau \to \infty$, $I_\nu \to S_\nu$.
Moreover, if $I_\nu > S_\nu$ holds, then $dI_\nu/d\tau_\nu < 0$ and I_ν decreases with increasing τ_ν. If $I_\nu < S_\nu$ holds, then $dI_\nu/d\tau_\nu > 0$ and I_ν increases with increasing τ_ν.

In words: with increasing optical depth τ_ν the intensity I_ν will always approach the value S_ν. S_ν is a property of the medium, while I_ν is determined by how much we can see of that medium.

2.4 Complication: Scattered Light

In the real world, things are further complicated by the scattering of light caused by dust particles, molecules, electrons, etc. An example is the so-called Rayleigh scattering, which is proportional to λ^{-4}! On Earth we know scattering in different forms. Molecules in the atmosphere Rayleigh-scatter sunlight, i.e especially blue light. By *multiple* scattering, blue sunlight reaches us from all directions causing the sky to appear blue. The strong reflection of car headlights on a dense fog is likewise caused by (back)scatter. In such a dense fog small drops of water cause a particularly strong scattering of light. A (deep-)red fog-light on a car can be seen at great distance, but the blue ANWB road-signs are almost indistinguishable.

In general, scattering means a loss of radiation intensity. In astronomy the combined effect of scattering and absorption is called *extinction*. The equation of radiative transfer can be extended by a scattering coefficient σ_ν; in its general form this leads to an equation that can not be solved analytically. Solutions are obtained by making simplifications, or by numerical calculations.

Chapter *3*

Blackbodies and Planck's Function

3.1 The Planck Function

(NB) Part of the contents of this Chapter are treated in more detail in *Radiative Processes in Astrophysics*, by G.B. Rybicki and A.P. Lightman (Ed. John Wiley and Sons, NY).

Thermal radiation is emitted by a medium which is in thermal equilibrium. *Blackbody radiation* is, in addition to this, in thermal equilibrium *with itself*. At any frequency, the intensity I of blackbody radiation is determined only by the temperature. It does not depend on the shape, nature or composition of the radiation source. We define:

$$I_\nu = f\left(\nu, T\right) = B_\nu\left(T\right) \ ,$$

where $B_\nu\left(T\right)$ is called the Planck function.

for blackbody radiation: $I_\nu = B_\nu$
for thermal radiation: $S_\nu = B_\nu$
so that: $j_\nu = \alpha_\nu \, B_\nu\left(T\right)$.
This is Kirchoff's law, which states that the more readily a body absorbs radiation the

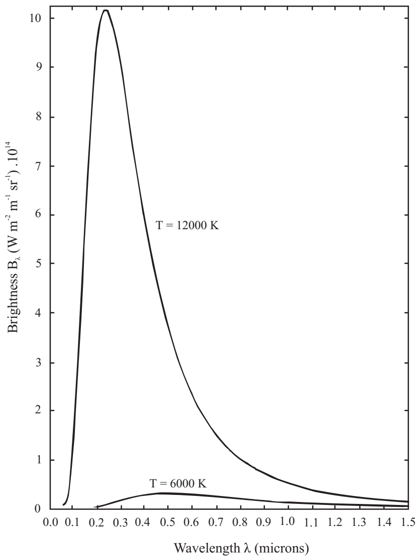

Figure 3.1 The Planck Function on linear scales, with brightness $B_\lambda\left(T\right)$ plotted as a function of wavelength λ for temperatures $T = 6000$ K (the Sun) and $T = 12000$ K. From: J.D. Kraus, *Radio astronomy*, 1966 McGraw-Hill, New York

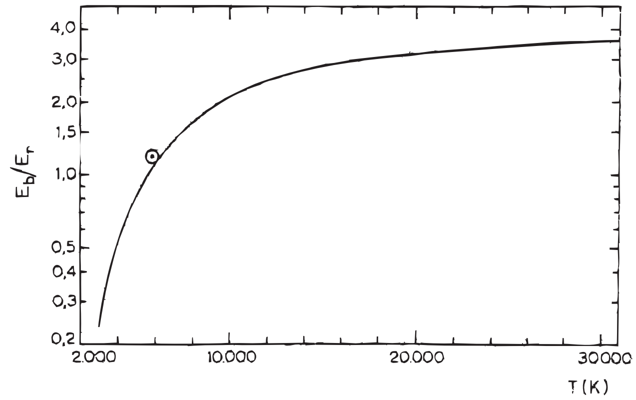

Figure 3.2 The 'colour' of celestial bodies is defined as the ratio of the fluxes or energies in two wavelength bands $\lambda_1 \pm \Delta\Lambda_1$ and $\lambda_2 \pm \Delta\Lambda_2$. The curve shows the energy ratio in photometric bands B ($\lambda_1 = 4400$ Å; $\Delta\Lambda_1 = 980$ Å) and R ($\lambda_2 = 7000$ Å; $\Delta\Lambda_2 = 2200$ Å) as a function of temperature T. The position of the Sun is marked with the symbol $\odot$. The colour itself is expressed as $(B - R) = -2.5 \log E_b/E_r$ (magnitude scale). From: J.van der Rijst, C. Zwaan *Astrofysica*, 1975

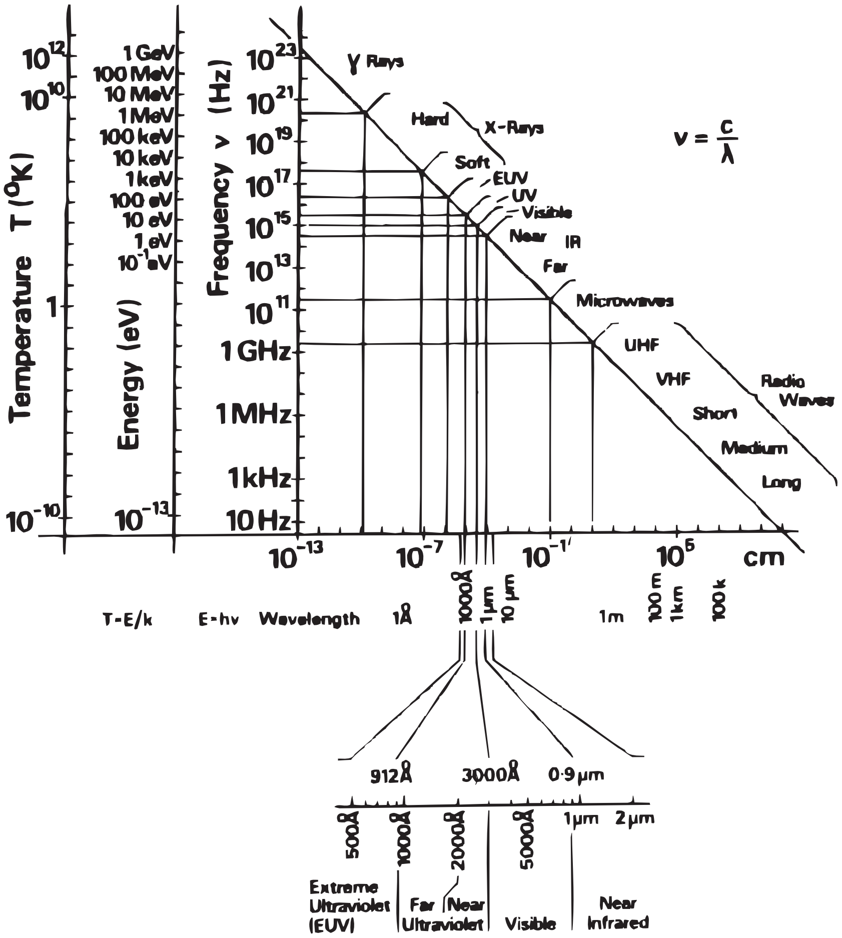

Figure 3.3 Relations between wavelength λ, frequency ν, energy eV and temperature T defined by the Planck Function. All scales are logarithmic.

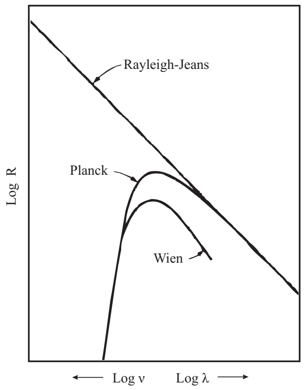

Figure 3.4 The Planck Function (logarithmic scale) coincides at long wavelengths with the Rayleigh-Jeans curve, and at short wavelengths with the Wien curve. Near the Planck 'peak' none of the other two curves is a good approximation of the Planck curve: Rayleigh-Jeans predicts too much radiation, Wien too little. From: J.D. Kraus, *Radio astronomy*, 1966 McGraw-Hill, New York

more readily it emits radiation.

A body enclosed by a blackbody surface is also in thermal equilibrium: $I_\nu = S_\nu = B_\nu(T)$; so that (see Section 2.3) in this case τ_ν should be infinitely large. In other words: every body or every layer may produce thermal radiation, but *only* a dense body or a thick layer (τ infinite) will produce blackbody radiation.

With Kirchoff's Law, the equation of radiative transfer equation for a blackbody becomes:

$$dI_\nu/ds = -\alpha_\nu\, I_\nu + \alpha_\nu\, B_\nu(T) \ , \qquad \text{or :}$$

$$dI_\nu/d\tau_\nu = -I_\nu + B_\nu(T) \ . \tag{3.1}$$

With thermodynamics and quantum theory the shape of the function $B_\nu(T)$ can be derived directly (cf. physics courses such as *'The Feynman Lectures on Physics'*, R.P. Feynman, R.B. Leighton, M. Sands, 1963; Part 1, pp. 41-3 to 41-8):

$$B_\nu(T) = 2\,h\,\nu^3/c^2\left[1/\left(e^{h\nu/kT} - 1\right)\right] \ . \qquad \text{(Planck's Law)} \tag{3.2}$$

where h is the Planck constant ($h = 6.625 \times 10^{-34}$ J s) and k is the Boltzmann constant ($k = 1.38 \times 10^{-23}$ J K^{-1}). Note that the quantity $h\nu/kT = T_\circ/T$ is dimensionless, and that for any $T_1 > T_2$, *always* $B_\nu(T_1) > B_\nu(T_2)$.

3.2 Rayleigh-Jeans and Wien Approximations of the Planck Function

a. If $h\nu << kT$ (the 'red' side of the Planck curve), we use a Taylor approximation: $e^{h\nu/kT} \approx 1 + h\nu/kT$: so:

$$I_\nu(RJ) = 2\,(\nu^2/c^2)\,kT \ . \qquad \text{(Rayleigh – Jeans (RJ) Law)} \tag{3.3}$$

In the Rayleigh-Jeans region, blackbody intensity is directly proportional to temperature and frequency squared (inversely proportional to wavelength squared).

N.B.: Define 'colour' as the radiation intensity ratio at two different frequencies; then all objects in the RJ region have the same 'colour' regardless of temperature.

b. If $h\nu >> kT$ (the 'blue' side of the Planck curve) we may approximate: $e^{h\nu/kT} - 1 \approx e^{h\nu}/kT$:

$$I_\nu(W) = 2\,h\,(\nu^3/c^2)\,e^{-h\nu/kT} \ . \qquad \text{(Wien's law (W))} \tag{3.4}$$

Observed area	λ (μm)	ν (Hz)	$T_o/T = h\,\nu/k\,T$ 10000 K	100 K
ultraviolet	0.1	3×10^{15}	$14.4\,(W)$	$1440\,(W)$
visual	0.5	6×10^{14}	$2.9\,(P)$	$290\,(W)$
infrared	50	6×10^{12}	$0.03\,(RJ)$	$3\,(P)$
radio	5000	6×10^{9}	$0.00003\,(RJ)$	$0.003\,(RJ)$

Table 3.1 Regime: W = Wien, RJ = Rayleigh-Jeans, P = Planck.

In Table 3.1 we compare radiation from two blackbodies, one with $T = 10000$ K (such as a star slightly hotter than the Sun) and one with a temperature $T = 100$ K (such as a warm dust cloud in space). Except for very hot stars, objects observed in the ultraviolet always fall in the Wien region. Unless they are very cool, objects observed in the radio region or in the infrared always fall in the Rayleigh-Jeans region. If we seek to determine an object's temperature from brightness (colour) measurements we must avoid measurement in the Rayleigh-Jeans region.

3.3 Radiation Peak and Total Brightness

Most astronomical objects have a Planck curve peaking in the visual or infrared part of the spectrum. This peak is found by differentiating the function $B_\nu\,(T)$ with respect to ν:

$$\nu_{max}/T = 5.9 \times 10^{10}\ \mathrm{Hz\,K^{-1}}\ . \qquad \text{(Wien Displacement Law)}$$

or differentiating $B_\lambda\,(T)$ with respect to λ:

$$\lambda_{max}\,T = 0.0029\ \mathrm{m\,K}\ . \tag{3.5}$$

The frequency (or wavelength) of the Planck peak shifts linearly with temperature. For example: the Sun radiates at $T \approx 5800$ K (see Section 4.1). Its intensity thus peaks at about $0.5\ \mu\mathrm{m}$ (5000 Å, 'green' light). The Earth and everything on it radiates at $T \approx 300$ K; peak intensities are thus at a twenty times larger wavelength (10 μm; NIR). For this reason, IR observations require cooling of the detector by e.g. liquid helium to substantially lower temperatures and preferably also cooling of the telescope itself. The latter is feasible only in the vacuum of space (see Section sec:Spaceobservatories)

By integrating the intensity $I_\nu = B_\nu\,(T)$ over all frequencies of the spectrum , we obtain the total brightness B_T:

$$B_T = \sigma\,T^4\,, \qquad (\mathrm{Stefan-Boltzman's\ Law}) \tag{3.6}$$

where the Stefan-Boltzmann constant σ has a value of $5.67 \times 10^{-8}\ \mathrm{W\,m^{-2}\,K^{-4}}$. A limited increase in temperature thus results in a very large increase in brightness.

For objects that are not blackbodies, we use the Stefan-Boltzmann Law to define the *effective temperature* T_{eff}:

$$B_T = \sigma\,T_{\text{eff}}^4 = \int \alpha_\nu\,B_\nu\,d\nu. \tag{3.7}$$

$0 < \alpha_\nu < 1$. Such an object is called a 'graybody' for which always: $T_{\text{eff}} < T_{BB}$!

Chapter 4

Planck and the Solar System

4.1 Temperature of the Sun

Even with a simple calorimeter (a glass of water with a thermometer!) and a relatively simple correction for the atmospheric absorption one can determine that at the orbit of the Earth the Solar flux (also known as the *Solar Constant* SC) amounts to:

$$\mathrm{SC} = 1.39 \times 10^3 \ \mathrm{W\,m^{-2}} \ .$$

According to the energy conservation law (Eqn. 2.1) and the Stefan-Boltzmann law (Eqn. 3.7) the following expression holds:

$$\mathrm{SC} \times (R/r)^2_{\mathrm{Sun}} = \sigma \, T^4_{\mathrm{eff}} \ ,$$

where R_{Sun} is the distance Earth-Sun (commonly referred to as the *Astronomical Unit* AU $= 1.5 \times 10^{11}$ m), and where $\mathrm{r_{Sun}}$ is the radius of the Sun. The ratio $r/R = 4.35 \times 10{-}3$ may be determined from a simple measurement of Sun's angular size. We find:

$$\sigma \, T^4_{\mathrm{eff}} = 6 \times 10^7 \ \mathrm{W\,m^{-2}} \ ,$$

implying: $T_{\mathrm{eff}}\,(\mathrm{Sun}) = 5800$ K.

Under low pressure, no known material is solid or liquid at this temperature. The Sun must therefore be gaseous. Note that for all of the above, no knowledge of the actual distance of the Sun is required.

Accurate determination of the Solar spectrum shows that the Sun does not radiate as a blackbody. This has a number of causes.

In the UV and the visual below approx. $5000\,\text{Å}(0.5\ \mu\mathrm{m})$ the Solar spectrum has a much lower intensity caused by the presence of many strong absorption lines.

Moreover:

- At every frequency or wavelength most of the radiation observed from the Sun originates at depths corresponding to $\tau \approx 1$. Radiation from greater depths orginates in non-transparent highτ layers and cannot be observed, while shallower layers contribute relatively little to the total radiation (cf. the equation of radiative transfer!).
- The absorption coefficient α_ν varies with frequency ν. In the red part of the spectrum α_ν is larger than in the blue part. As a result, in the red $\tau_\nu = \int \alpha_\nu \mathrm{d}s = 1$ is reached at a shallower depth ds into the Sun than in the blue.
- Because Solar temperatures increase rapidly with increasing depth, red light region originates in cooler layers than blue light. As a result, the solar spectrum cannot be characterized by one single temperature, which is a requirement for a blackbody.

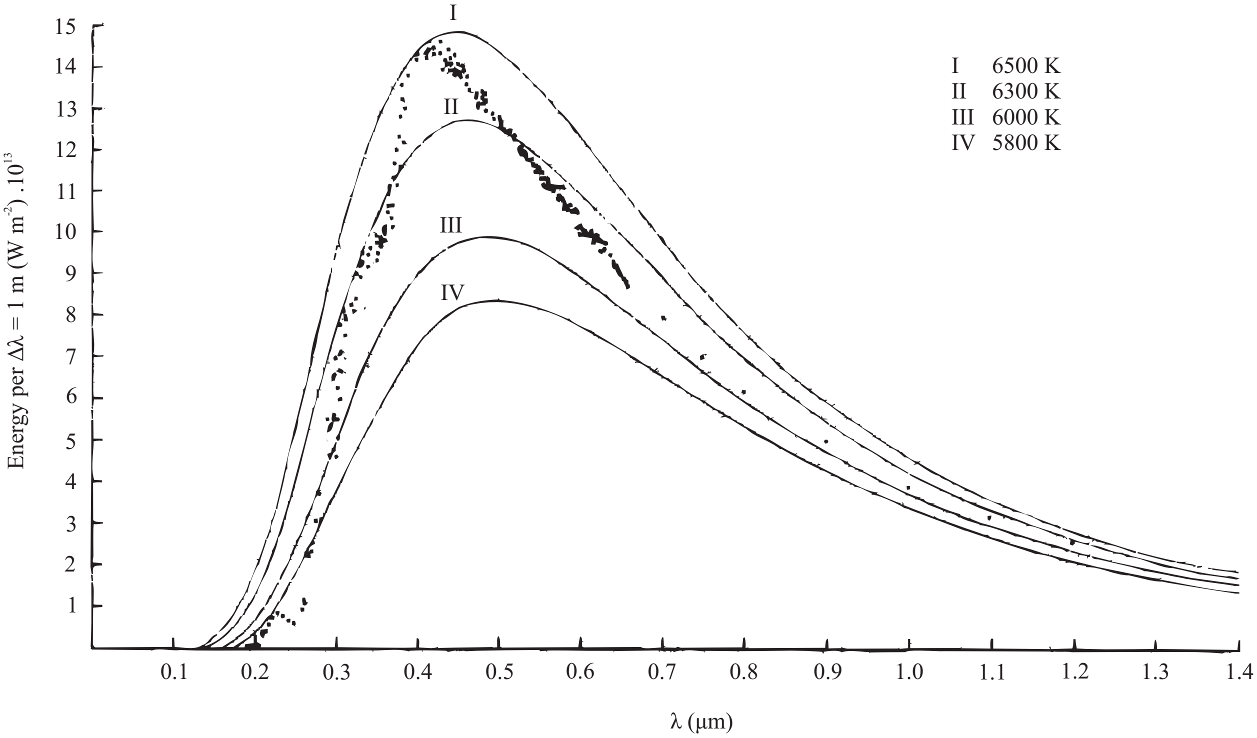

Figure 4.1 Actual radiation curve at the center of the Solar disk (points), compared to Planck functions for temperatures between $5800\ \mathrm{K}$ and $6500\ \mathrm{K}$. From: D. Labs & H. Neckel, 1972 in *Solar Physics* 22, 64

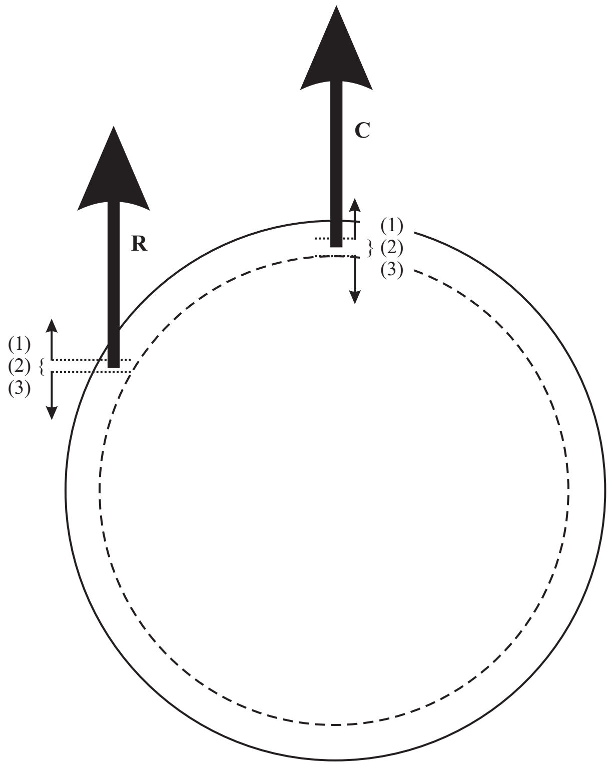

Figure 4.2 For observations at the edge (R) the radiating layer (2) is located further away from the center than for observations at the center (C) of the disk. In this Figure, (1) corresponds to layers which are almost wholly transparent ($\tau << 1$), while (3) corresponds to non-transparent layers ($\tau >> 1$). From: J. van der Rijst & C. Zwaan *Astrofysica*

More general: at every frequency the Solar radiation originates in gas at a different depth and thus a different temperature. For this reason, the temperature assigned to the Solar 'surface' depends on the frequency or wavelength of observation. For example: at $4250\,\text{Å}(0.425\ \mu\text{m})$ the layer with $\tau = 1$ has $T = 6500$ K and at $8500\,\text{Å}(0.85\ \mu\text{m})$ it is at $T = 6100$ K. At wavelengths $> 1\ \mu\text{m}$ the temperature of the 'visible surface' is indeed 5800 K. At all other wavelengths temperatures are between these values.

The Solar disk is brighter in the center than at the edge (limb darkening). This is caused by the fact that at the edge, we view the Solar atmosphere almost side-on. Because of the larger pathlength through the atmosphere, optical depths $\tau \approx 1$ are reached further from the center of the Sun (i.e. at lower geometrical depths) than is the case if one looks at the center of the Solar disk. Planck's law predicts that limb darkening will be more pronounced in blue than in red light.

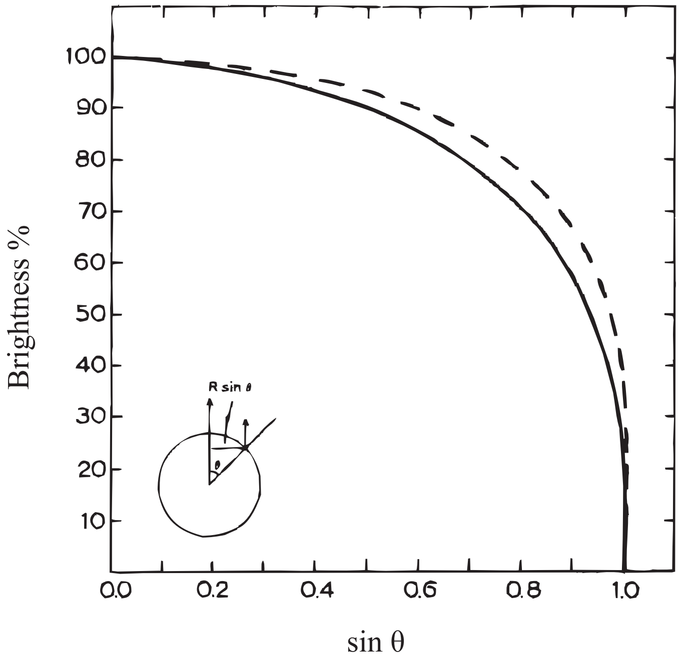

Figure 4.3 Relative brightness across the Solar disk (center $= 100\%$). The solid line corresponds to blue light (4500 Å) and the dashed line corresponds to orange-yellow light (6000 Å). From: J. van der Rijst & C. Zwaan *Astrofysica*

4.2 Surface Temperatures of the Planets

The Stefan-Boltzmann law can also be used to calculate the effective temperatures of the planets.

The energy flow leaving the Solar surface each second is:

$$\begin{aligned} E_{\text{Sun}} &= 4\,\pi\, r_{\text{Sun}}^2 \;\times F_{\text{Sun}} \\ &= 4\,\pi\, r_{\text{Sun}}^2 \;\times B_{T,\,\text{Sun}} \\ &= 4\,\pi\, r_{\text{Sun}}^2\, \sigma\, T_{\text{eff}}^4\,, \end{aligned}$$

where T_{eff} is the effective temperature of the Sun and r_{Sun} is the radius of the Sun. At a distance R_{planet} therefore (energy conservation; see 2.1):

$$F_{\text{planet}} = E_{\text{Sun}} / \left(4\,\pi\, R_{\text{planet}}^2\right) \quad = \left(r_{\text{Sun}}/R_{\text{planet}}\right)^2 \sigma\, T_{\text{eff}}^4 .$$

In general planets are no blackbodies, but absorb only a fraction $(1 - a)$ of the incoming radiation; the fraction a is reflected and is called the *albedo* of the planet. Clouds have a high albedo ($a = 0.7$), while porous (volcanic) rock has a low albedo ($a = 0.1$). Define:

$$B_{\mathrm{abs}} = (1 - a)\, F_{\mathrm{planet}} \,.$$

In temperature equilibrium, the energy absorbed must be balanced by the energy radiated by the planet, so that:

$$B_{\mathrm{planet}} = \sigma\, T^4_{\mathrm{planet}} = B_{\mathrm{abs}} \,,$$

where T_{planet} is the *effective mean* temperature of the subsolar point on the planet. So:

$$\sigma\, T^4_{\mathrm{planet}} = \sigma\, T^4_{\mathrm{eff}}\, (r_{\mathrm{Sun}}/R_{\mathrm{planet}})^2\, (1 - a) \ ,$$

This assumes that the absorbed Solar energy transported across the planet's surface (planets are poor conductors) and that, as a result the absorbed energy is reradiated locally. This is, however, not correct if the planet rotates rapidly and, in particular, if the planet has an atmosphere sufficiently dense to effectively transport heat by convection (winds). In such cases, differences between day and night temperatures are relatively small, as are latitudinal differences. In such cases, the Solar energy is *received* over a projected surface $\pi\, r^2_{\mathrm{planet}}$ but the absorbed energy is *emitted* over the larger spherical surface $4\,\pi\, r^2_{\mathrm{planet}}$ thereby reducing average temperatures by a factor $4^{-1/4} = 0.7$, so that we obtain:

$$T_{\mathrm{planet}} = 0.7\ T_{\mathrm{eff}}\, (r_{\mathrm{Sun}}/R_{\mathrm{planet}})^{1/2}\, (1 - a)^{1/4} \tag{4.1}$$

Obviously, these results apply to *spacecraft* as well as planets, provided that the correct albedo is used! Such calculations, also taking into account the heat produced by on-board hardware, are of importance when dealing with the thermal control of spacecraft. A special application occurred during the manned lunar missions in the early seventies. During flight, the Apollo spacecraft were made to rotate, decreasing its temperature became (barbecue effect) just as happens with fast-rotating planets. Even lower spacecraft temperatures are obtained by passive cooling techniques such as, for instance, increasing the albedo of the absorbing surface while decreasing the albedo of the radiating surface, or by increasing the radiating surface area with respect to the absorbing surface area.

4.3 Planetary Temperatures: Theory and Reality

Because planets are generally at distances $R_{\mathrm{planet}} >> r_{\mathrm{Sun}}$, their temperatures are expected to be at most several hundred K. Thus, we should measure planetary temperatures in the infrared, where the radiation peak corresponding such temperatures occurs. In the visual region we see only reflected sunlight (albedo!) which is not a measure of plantary temperature.

In Table 4.1 we present for the planets of the Solar System distance to the Sun R_{plan} (in AU = 150 million km), albedo a, and the predicted and actual temperatures (in Kelvins). Note that the calculation of the predicted values only require *relative* distances, as e.g. provided by Kepler's age-old third law.

Mercury, Mars and Pluto have a low albedo (rock). The other planets have high albedos (clouds, and ice in the case of Pluto). For Mercury and Pluto, the predicted temperatures are effectively identical to those actually observed value, but for all other planets Venus the observed exceed the predicted values.

The observed temperatures of Jupiter, Saturn, Uranus and Neptune correspond to the top of the cloud layer (1 bar pressure level). These planets radiate additional energy from their interiors (as does the Earth to a minor degree). This energy appears to result from chemo-physical processes deep in the gaseous interior.

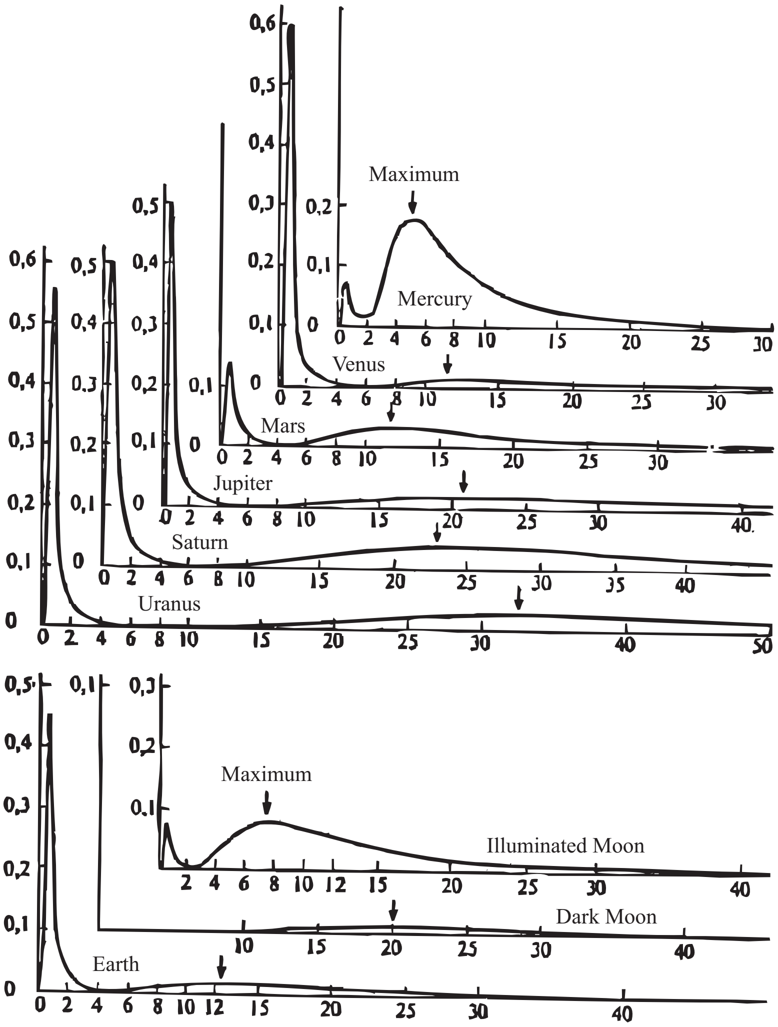

Figure 4.4 Radiation from the Moon and the Planets. The wavelength is marked in microns ($1\ \mu\mathrm{m} = 10^{-6}\ \mathrm{m}$). From M. Minnaert & J. Houtgast, 1946 *Hemel en Dampkring*

Planet	R_{planet} (AU)	a	Temperature Predicted (K)	Observed (K)
Mercury	0.39	0.11	430	440
Venus	0.72	0.65	240	730
Earth	1.00	0.37	250	290
Mars	1.52	0.15	215	240
Jupiter	5.20	0.52	110	165
Saturn	9.54	0.47	85	135
Uranus	19.2	0.51	55	76
Neptune	30.1	0.41	50	73
Pluto	39.5	0.5	55	58

Table 4.1 The planets' distance to the Sun, their albedo, and their calculated and observed temperatures.

4.4 Greenhouse Effect

The temperature excess of the Earth, Venus and Mars, however, must be explained differently. Venus and Mars have thick respectively thin atmospheres consisting almost entirely of carbon dioxide. The Earth's atmosphere contains a small amount of carbon dioxide, but also water vapour, methane and nitric oxides. These presence of these greenhouse gases causes temperatures to be higher than predicted (greenhouse effect).

The fraction of incident solar energy ($T_{\rm eff} = 5800$ K, radiation peak at visual wavelengths) that is not reflected reaches the planetary surface and lower atmosphere essentially unhampered, heating them to a temperature $T_{\rm planet}$. In an equilibrium state, all absorbed energy must be reradiated at $T_{\rm planet}$. This temperature is much lower than $T_{\rm eff}$ and reradiation therefore occurs at the much longer wavelengths of the near-IR (Wien's displacement law). However, at *these* wavelengths strong absorption by greenhouse gas molecules renders the atmosphere much less transparent for the outgoing radiation than it was for the incoming Solar radiation. Due to the absorbed energy, the atmospheric temperature is increased. At higher temperatures, emission (outside the absorption bands) is intensified, and this compensates for the decreased efficiency of reradiation caused by the greenhouse absorption. N.B. Although the *actual temperature* of a planet is influenced by the presence of greenhouse gases effect, the *effective temperature* of a planet is the same irrespective whether or not a greenhouse effect exists!
Of the three 'terrestrial' planets only the Earth itself is suited to sustain human live. Mars has a thin atmosphere and is too cold in spite of its greenhouse effect.. Venus has a very dense atmosphere and is much too hot. Both extreme 'climates' are quite stable, giving raise to questions like: How stable is the Earth's climate? Can a 'runaway greenhouse effect' occur on Earth? What does it take to induce an eternal ice age? Answers require better atmospheric models, demonstrating the importance of using spacecraft to study the climate on neighbouring planets and to carefully monitor our own atmosphere ('Mission to Planet Earth').

4.5 Scale of the Solar System

Until now, we have determined fundamental properties without knowing the absolute distances of the Sun and the planets. To make further progress, however, we have to change this.

Relative distances of all objects orbiting the Sun follow from Newton's law and from Kepler's Third Law in particular.

$$F = G\,m\,M/r^2\,, \qquad \text{(Newton's law)}$$

whith the gravitational constant $G = 6.67 \times 10^{-11}$ m^3 kg^{-1} s^{-2} (*do not memorize!*).

In addition:

- Acceleration in a circular orbit: $a = v^2/r$;
- Period in a circular orbit: $T = 2\,\pi\,r/v \;\rightarrow\; v^2 = 4\,\pi^2\,(r/T)^2$.

  Thus:

$$4\,\pi^2\,R/T^2 = G\,M/R^2\,,$$

from which follows:

$$R^3/T^2 = \left(G/4\,\pi^2\right)\,M\,. \qquad \text{(Kepler's third law)}$$

N.B.: this derivation is only valid if $m << M$; in general this expression is:

$$R^3/T^2 = \left(G/4\,\pi^2\right)(m+M)\,. \tag{4.2}$$

The planetary motions in the Solar System are dominated by mass M of the Sun, all planets having masses $m << M$. The most massive of all, Jupiter, does not rise above $m_{\text{Jupiter}} = 0.001\, M_{\text{Sun}}$. Kepler's Third Law defines a *scaled model* of the solar system using only easily determined orbital periods. Measurement of a single absolute distance anywhere in this model determines the scale factor and all other absolute distances.

The distance of the Moon can be determined trigonometrically, but is irrelevant, as the Moon orbits the Earth, not the Sun. The planets are too distant for succesful trigonometry.

In the 18^{th} century, the first reasonably accurate determination of the scale of the Solar system was derived from observations of the 1769 transit of Venus across the Solar disk (i.e. a line-up of the Earth, Venus and the Sun) both in Greenwich and on Tahiti (captain James Cook). Using Kepler's model, the size of the Earth – already well-known – and accurate time measurements the absolute distance to Venus was determined (*Can you work out how they did it?*).

The Solar System also contains smaller bodies, asteroids, orbiting the Sun in often highly excentric ellipses. One of these, Eros, passed by so closely in 1931 that its distance to the Earth could be determined trigonometrically, providing a new and more accurate scale factor!

Nowadays, the scaling factor usually expressed as the Astronomical Unit AU (i.e. the mean distance of the Earth to the Sun) has been determined with very high accuracy from the orbital mechanics of interplanetary spacecraft, and from planetary radar experiments, especially those directed towards Venus.

The result: $1\ \text{AU} = 149597871\ \text{km}$, is usually rounded off to 150 million kilometers.

4.6 Mass, Luminosity and Radius of the Sun

Knowing the Sun's distance $R(= 1\ \text{AU})$ and the Solar Constant SC, we easily calculate the luminosity of the Sun, i.e. the total power radiated by the Sun per unit time:

$$L_{\text{Sun}} = 4\,\pi\,R^2\,SC \;=\; 4 \times 10^{26}\ \text{W}\ .$$

Knowing the absolute value of the Astronomical Unit and the Earth's orbital period T = 1 year = 3.16 $\times 10^7$ sec we equally easily calculate the mass of the Sun, using Kepler's Third Law:

$$M_{\text{Sun}} = 2 \times 10^{30}\ \text{kg}\ .$$

Compare: $m_{\text{Earth}} \;=\; 6 \;\times\; 10^{24}\ \text{kg} \;=\; 3 \;\times\; 10^{-6}\,M_{\text{Sun}}$, which is calculated in a very similar manner from the motion of the Moon, or indeed any (artificial) satellite orbiting the Earth. In fact, for all planets with moons the mass can be readily calculated. This is more difficult for planets lacking moons, e.g. Mercury and Venus. The masses of these planets used to be determined in a rather laborious manner from the disturbances they caused in the orbit of the Earth, and in each other's orbit. Now, of course, much more accurate mass determinations have been yielded by the motion of spacecraft near these planets. This is also true for our own Moon. *Question: how did we know the Moon's mass before the era of space travel?*

The radius of the Sun follows from the magnitude of the Astronomical Unit and its mean angular diameter as seen from the Earth (θ_{Sun} = 32 arcminutes = 9.3 $\times$ 10^{-3} rad):

$$r_{\text{Sun}} = 1/2\;\theta\,R \;= 6.96 \times 10^{8}\ \text{m}\ .$$

Compare: $r_{\text{Earth}} = 6.37 \times 10^6\ \text{m} = 0.9 \times 10^{-2} r_{\text{Sun}}$.

Finally, we may answer the question why we see a sharp Solar image, even though we are dealing with a gaseous body which, in reality, does not have a sharp outline. Going deeper into the Sun, pressure rapidly increases. In turn, the mounting pressure causes the density, and with it the absorption coefficient α_ν to increase equally rapidly. Because of this rapid increase in α_ν the layer corresponding to $\tau \approx 1$ is very thin: approx. $100 \text{ km} = 10^5 \text{ m}$, or 0.01 per cent of the diameter of the Sun. This explains why the outline of the Sun can be seen sharply: the thickness of the transition layer is about 0.14 arcseconds, i.e. four hundred times smaller than the resolving power of the human eye.

Chapter 5
Thermal Radiation

5.1 Thermal Processes

Every object in the universe with a temperature above absolute zero (i.e. *everything*) emits radiation. If the radiative intensity I is a function $f(T)$ determined by the temperature T, one speaks of *thermal radiation*; if it is not, of non-thermal radiation (see Appendix B). In general, thermal radiation of a gas finds its origin in changes of the energy states of the particles in the gas. For instance, a negatively charged particle (electron) has a certain amount of energy with respect to a positively charged particle (ion) dependent on the distance between electron and ion. If the electron is bound to the ion, this energy can only have discrete values: the system is quantized. If the electron is 'free', its energy with respect to the ion can in principle have any (positive) value. We distinguish:

Free-Free Transitions: a moving free electron is deflected by an ion. It loses energy which is released in the form of a photon. This can, in principle, possess any amount of energy. In a gas, at any time a large number of electrons and ions interact, causing this process to provide radiation, in principle at every conceivable wavelength (*continuum radiation*).

Free-Bound Transitions: a free electron is captured by an ion into one of its discrete energy levels. The energy difference, released as a photon, may have any value above *at least* the binding energy of the ion at that level (equal to the ionization energy). Continuum radiation results, but only at wavelengths shorter than that corresponding to the ionization limit.

Bound-Bound Transitions: a bound electron transits from one energy level to a lower energy (less excited) level. As these energy levels may have only specific (discrete) values, radiation can occur only at particular, ion-specific, frequencies or wavelengths (*line radiation*).

The last two processes also work in reverse, requiring energy input. This may, for instance, be provided by the absorption of a photon (photo-excitation) or by a collision with another particle (collisional excitation).

In the following, we will take a closer look at the manner in which these spectra are produced.

5.2 The Bohr Atom

A simple but in many cases satisfactory model of atomic structure was proposed by Niels Bohr in the early 20^{th} century. An electron with momentum p_e orbits an atomic nucleus. Following de Broglie, associated with this electron, as indeed with every particle, is a probability wave of wavelength

$$\lambda_e = h/p_e = h/m_e v. \tag{5.1}$$

The path of the electron around the nucleus cannot be of any size, but must be an integral number times this wavelength:

$$2\pi r_n = n\,\lambda_e. \qquad (n = 1, 2, 3...)$$

In a stable electron orbit, the centrifugal effect is balanced by the attractive (electromagnetic) force:

$$m_e v^2/r = e^2/r^2 \rightarrow m_e^2 v^2 = m_e e^2/r$$

Note that for convenience (cf. Feynman Lectures on Physics, Vol. I, p. 12.7) we have defined $e^2 = q^2/4\pi\epsilon_o$, with q the charge on an electron. Combine all this and find:

$$r_n = n^2\,[(h/2\pi)^2\; 1/(m_e e^2)] \tag{5.2}$$

The expression between square brackets, i.e. the electron orbit radius corresponding to $n = 1$, is called the Bohr radius.
An electron orbiting at radius r_n has a total energy

$$\begin{aligned} E_n &= E(kin) + E(e) \\ &= 1/2\,(m_e v^2) - e^2/r \qquad \text{(negative energy!)} \\ &= -1/2\;(1/n^2)\;(m_e\,e^4)\;(2\pi/h)^2. \end{aligned}$$

If $n \rightarrow \infty$, $E \rightarrow 0$: an electron *just unbound* has zero energy. However, to unbind an electron, energy must be provided!

The transition of an electron from a higher energy (excited) state m to a lower energy (less excited) state n, i.e. $m > n$ releases energy:

$$\begin{aligned} \Delta E &= E_m - E_n \\ &= 1/2\;\left(1/n^2 - 1/m^2\right)\;(m_e\,e^4)\;(2\pi/h)^2 \end{aligned} \tag{5.3}$$

as a foton at a wavelength $\lambda = hc\,/\,\Delta E$.

5.3 The Hydrogen Line Spectrum

Spectral lines arise exclusively from such bound-bound transitions. We now consider the line spectrum of the simplest element, hydrogen. The wavelengths or frequencies of the lines occurring in that spectrum are written as:

$$\nu_{nm}/c = R\left(1/n^2 - 1/m^2\right)\,. \qquad (n < m)$$

For hydrogen (H) the so-called Rydberg constant is $R = 109667.576\ \mathrm{cm}^{-1}$. Each level n produces its own series of lines. As m increases, the sepration of lines decreases; past $1/m^2 = 0$ a free-bound continuum occurs.

Astronomical nomenclature:
$n = 1$: Lyman series;
$n = 2$: Balmer series;

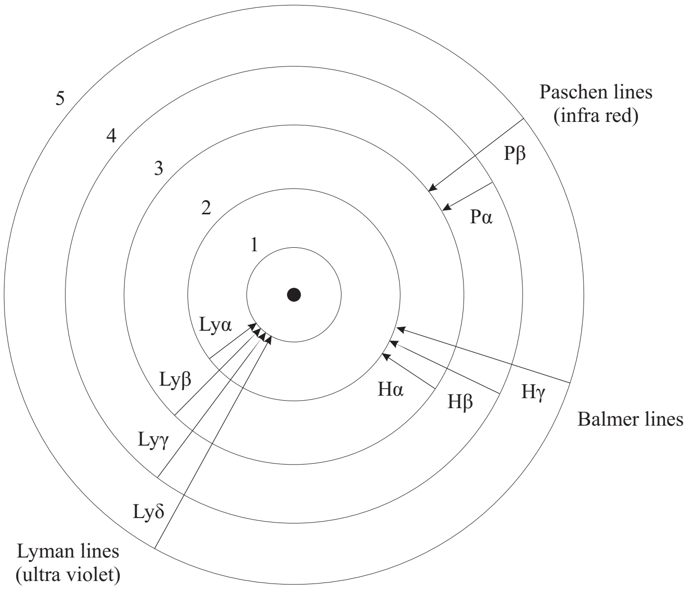

Figure 5.1 Quantum orbits in the hydrogen atom and the origin of spectral line series. From: J. van der Rijst & C. Zwaan *Astrofysica*

$n = 3$: Paschen series;
$n = 4$: Brackett series.

Transitions $m \rightarrow m - 1$ are indicated as α, $m \rightarrow m - 2$ as β, etc.

Examples:

$$H\alpha = \mathrm{Balmer} - \alpha = H(3-2)\,;\ \mathrm{Br} - \gamma = \mathrm{Brackett}\,\gamma = H(7-4)\,.$$

Assignment: Calculate in which wavelength regions the Lyman, Balmer and Brackett series are located.

N.B.: The energy of a photon is $E = h\,\nu$. The Rydberg equation therefore also indicates the energy levels corresponding to spectral lines. These energy levels are often expressed in eV (electron-volt): $1\ \mathrm{eV} = 1.60 \times 10^{-12}\ \mathrm{erg} = 1.6 \times 10^{-19}\ \mathrm{W\,s}$.

Assignment: Calculate the ionization limit of the Lyman series in eV*. What is the corresponding wavelength?*

For *hydrogen-like* elements (helium^{+}, lithium^{++}) the Rydberg constant R is replaced by R', where $R' = Z^2\,R$; Z is the atomic number. For other atoms/ions the calculation of the energy levels is more difficult and our knowledge is mainly empirical. This is even more so in the case of molecules.

It is possible to obtain information on temperature and particle density by comparing the intensities of well-chosen spectral lines.

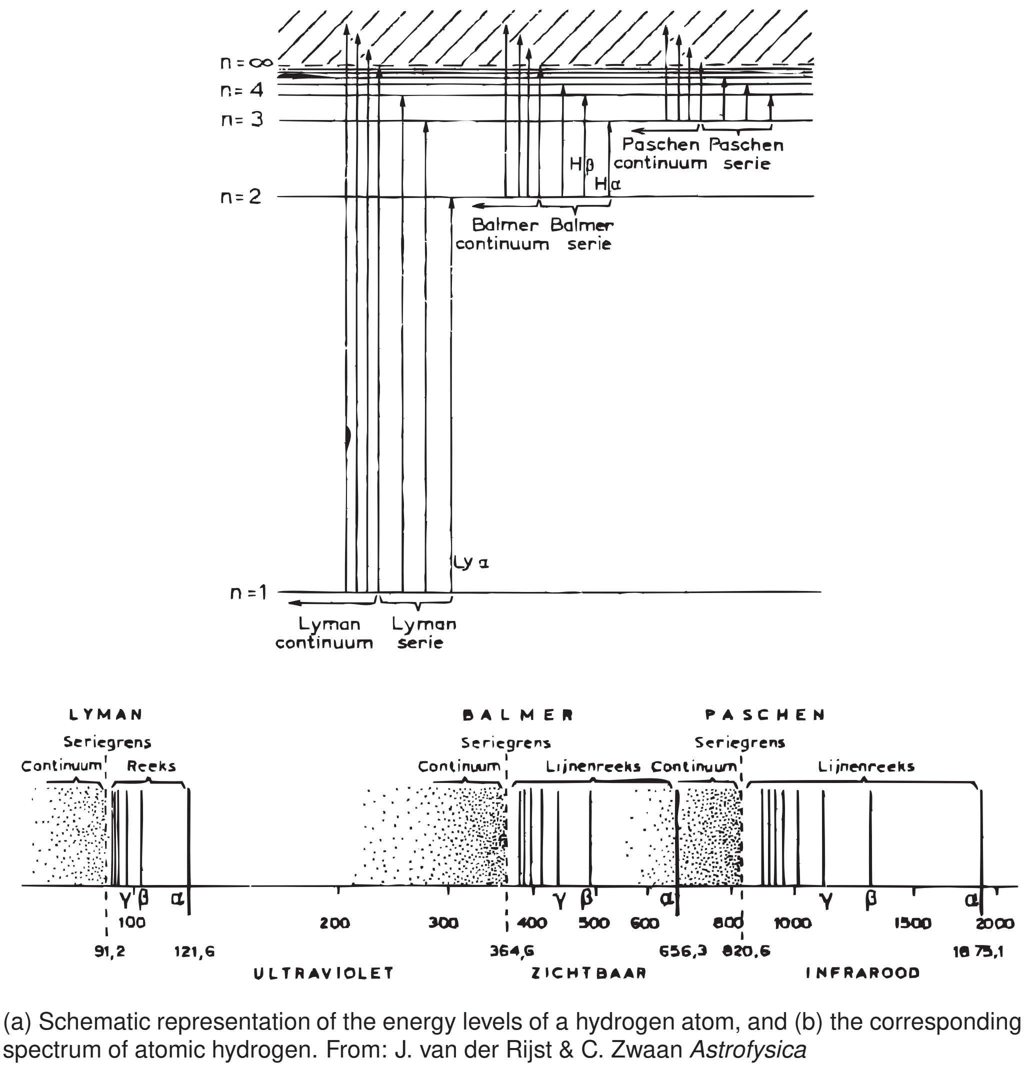

Figure 5.2 (a) Schematic representation of the energy levels of a hydrogen atom, and (b) the corresponding spectrum of atomic hydrogen. From: J. van der Rijst & C. Zwaan *Astrofysica*

Chapter 6

Saha's Law and Spectroscopic Analysis

6.1 Ionization Equilibrium

(Note: much of this Chapter and the next is based on the text *Astrofysica* by J. van der Rijst and C. Zwaan, prepared in 1975 for the Commissie Modernisering Leerplan Natuurkunde)
Within stars such as the Sun, pressure, density and temperature all rapidly increase with depth. Before the structure and evolution of stars can be studied, we first have to determine their composition. Spectroscopy provides the required information. The radiation spectra of gases consist of lines and continuum. The free-bound continuum starts past the wavelength corresponding to the ionization energy of the atoms making up the gas. *Below* this ionization limit, atoms can only absorb energy at discrete levels (lines); *above* this limit it can absorb any amount of energy (continuum). Every element has its own characteristic sets of lines.

The degree of ionization of an element will increase with increasing temperature, but there is an important complication. In excited atoms, electrons spontaneously fall back to lower energy levels. However, an ionized atom (ion: an atom which has lost one or more electrons) may become a neutral atom (whether or not excited) by absorbing one or more free electrons. Thus, the greater electron-densities N_e (the number of electrons per cm^3) are, the greater the probability of recombination is, and consequently the lower the number of ions will be).
For every ionization process, there is a pressure equilibrium equation of the form:

$$x = [p\,(\mathrm{ion}) \times p\,(e)] \,/ [p\,(\mathrm{atom})] = f\,(T, E_i) \ .$$

As partial pressures are proportional to particle densities, we may replace p by N:

$$x = [N\,(\mathrm{ion}) \times N_e] \,/ [N\,(\mathrm{atom})] = N_i/N_a \times N_e = f\,(T, E_i) \ .$$

Apply the Boltzmann distribution law to find $f\,(T, E_i)$:

$$x = N_i/N_a \times N_e = \mathrm{cst}\ T^{3/2}\ e^{-E_i/kT} \ ,$$

where the constant is given by $(2\,\pi\,m_e\,k)^{3/2}/h^3$. However, a star contains not just a single ionized element but many others as well. Electrons released by these other elements also contribute to the ionization equilibrium of the element under consideration. For this reason we reintroduce a partial pressure for electrons, this time for *all* electrons:

$$N_i/N_a = \mathrm{cst}\,(T^{5/2}/p_e)\ e^{-E_i/kT} \ . \tag{6.1}$$

This is known as Saha's law. *Note the extra factor T introduced by the ideal gas law!*

Saha's law provides the ratio between the numbers of atoms and ions in a gas as a function of E_i and T. The Boltzmann distribution law shows how each of these atoms and ions is distributed over the different energy levels as a function of T.

6.2 Absorption Line Analysis

For each atom or ion, the number of particles per unit surface area or column density in a particular ionization state and a particular energy level can be determined from the intensity of the corresponding spectral line and its Einstein transition probabilities (see Appendix A). By measuring spectral line intensities for different elements in various ionization states and energy levels, one may eventually determine the relative numbers of the different elements and, as a result, the composition of a radiating gas.

Temperature T	n_2/n_1	n_2/n_1	n_2/n_1
3000 K	$2.8\ 10^{-17}$	$4.3\ 10^{-20}$	$2.3\ 10^{-21}$
5000 K	$2.1\ 10^{-10}$	$5.8\ 10^{-12}$	$2.6\ 10^{-12}$
10000 K	$2.9\ 10^{-5}$	$7.2\ 10^{-6}$	$1.6\ 10^{-5}$
20000 K	$1.1\ 10^{-2}$	$8.0\ 10^{-3}$	$4.0\ 10^{-2}$
40000 K	$2.0\ 10^{-1}$	$3.0\ 10^{-1}$	$2.0\ 10^{-0}$

Table 6.1 Distribution of atoms over various energy levels, calculated from Boltzman's distribution law. Note that in stars like the Sun, almost all atoms are in the ground state.

Temperature	Electron Pressure p_e			
T	0.1 N/m^2	1 N/m^2	10 N/m^2	100 N/m^2
3000 K	$2.5\ 10^{-15}$	$2.5\ 10^{-16}$	$2.5\ 10^{-17}$	$2.5\ 10^{-18}$
5000 K	$1.2\ 10^{-5}$	$1.2\ 10^{-6}$	$1.2\ 10^{-7}$	$1.2\ 10^{-8}$
10000 K	$4.7\ 10^{2}$	$4.7\ 10^{1}$	$4.7\ 10^{1}$	$4.7\ 10^{0}$
20000 K	$7.1\ 10^{6}$	$7.1\ 10^{5}$	$7.1\ 10^{4}$	$7.1\ 10^{3}$
40000 K	$2.1\ 10^{9}$	$2.1\ 10^{8}$	$2.1\ 10^{7}$	$2.1\ 10^{6}$

Table 6.2 Ratios between the number of hydrogen ions and atoms for different temperatures and electron pressures occurring in stellar atmospheres. In stars like the Sun, hydrogen will be mostly atomic.

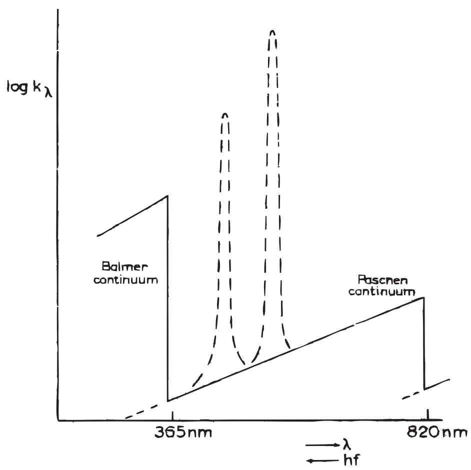

Figure 6.1 Very schematical behaviour of the absorption coefficient in hot atomic hydrogen. The continuous absorption coefficient is marked by a solid line, and the line absorption coefficients of two Balmer lines by dotted lines.

Stars have temperatures increasing strongly with depth. Observation of spectral lines arising from energy levels at different temperatures and with different opacities therefore allow us to probe stars not just at the surface, but out to some depth as well. In

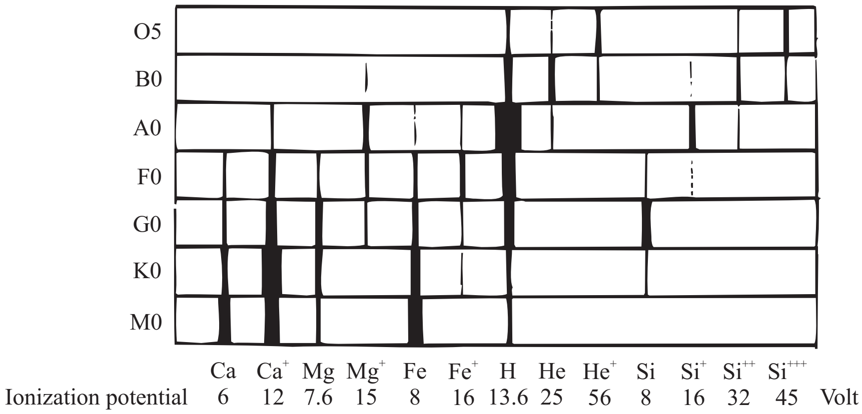

Figure 6.2 Relation between the ionization potential of some elements (atoms as well as ions) and the intensities of their absorption lines in the atmosphere of stars with increasing surface temperatures. Compare with Table 7.2

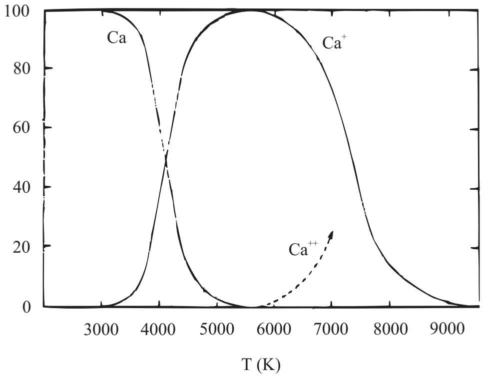

Figure 6.3 Ionization of calcium in stellar atmospheres. With increasing temperature, the neutral atoms are progressively converted into singly ionized atoms. Further temperature increases turn these into doubly ionized atoms. Almost no netral atoms are left, and the concentration of singly ionized atoms drops.

stars, spectral lines are usually observed in absorption. For stars that are not too hot ($T < 20000$ K) the absorption coefficient α_ν is much larger in the Balmer continuum than in the Paschen continuum (see Figure 6.1), because the number n_2 of particles in the second energy level is much larger than the number n_3 of particles in the third energy level. *What about the Lyman continuum? Find the answer yourself!*

The line absorption coefficient is always larger, in many cases very much larger by up to a factor of 10^6, than the continuous absorption coefficient at the same wavelength. However, in all lines of a particular series such as the Balmer, the line absorption coefficients are again proportional to the number n_2 etc. of particles in the lower level. The lines from such a series have different intensities, related to the transition probabilities (Appendix A): $A_{21} > A_{32} > A_{42}$ etc.

Stars with different temperatures consist of gas in different degrees of ionization and, therefore, will show the same absorption lines with different intensities. Radiation received from a star over a certain wavelength interval originates in a thin layer. The depth of the radiating layer in the star is inversely proportional to the absorption coefficient of the gas in that wavelength interval (continuum radiation originates much deeper into the star than line absorption); absorption lines of different elements and transitions will sample (sometimes very) different depths. By measuring spectral lines with decreasing absorption coefficients, we may deduce how temperature changes with increasing depth within the star;
By comparing absorption line intensities of *a particular element*, we may deduce the number of particles in the various states of ionization. By comparing absorption line

intensities of *different* elements, the relative number of particles (abundances) of these elements can also be determined.

Diligent work over the years has provided the following result for the Sun:

Element	Percentage by Number	Element	Percentage by Number
H	90	Mg	0.0038
He	10	Si	0.0035
C	0.036	S	0.0016
N	0.011	Fe	0.0035
O	0.085	the rest	0.04
Ne	0.012		

It turns out that almost all stars (except for the oldest ones) have more or less the same composition. The above ratios are therefore also called the 'cosmic abundance'.

Chapter 7

Properties of Stars

7.1 Spectral classification

Stellar spectra generally show many absorption lines superposed on a continuum. These absorption spectra can be classified according to continuum shape and particularly the occurrence and depth of various absorption lines. At the end of the 19^{th} century, spectra could be photographed but the classification was entirely empirical as physical insight was lacking. Eventually, the original alphabetical classification had to be revised and changed into the physically meaningful sequence O, B, A, F, G, K, M (with subclasses: B0, B1, B2, ..., B9, etc). Spectral class also specifies the colour of a star: from blue-white (O and B) via white, yellow (G) and orange to red (M). Initially it was thought that a difference in spectrum meant a difference in composition. However, in 1922, Russell used the Boltzmann and Saha equations to show that the revised spectral sequence is a temperature sequence (see Figure 7.1 and Table 7.1).

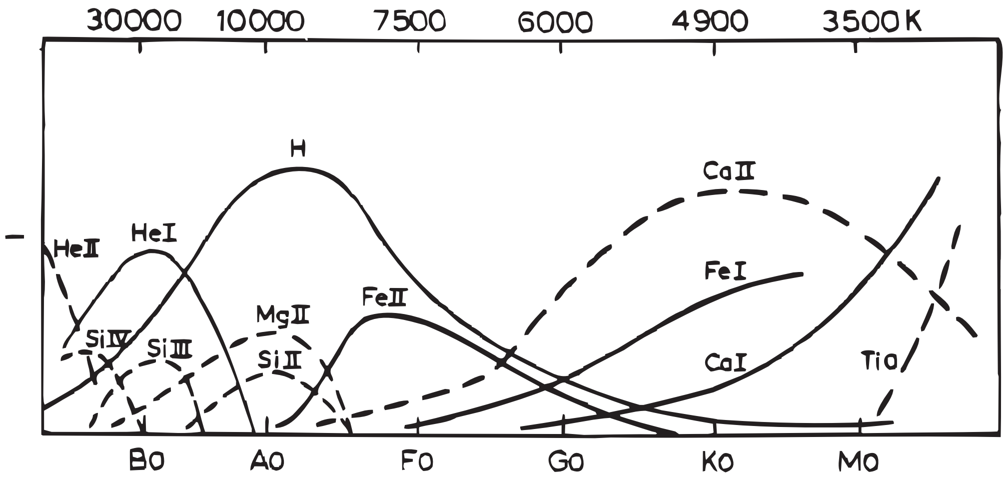

Figure 7.1 Relation between spectral line (absorption) strength and effective temperature of stars.

Spectral types occurring at the start of the revised sequence are called 'early' and belong to relatively hot stars. Those at the end of the sequence are called 'late' and belong to relatively cool stars. Early-type stars have spectra showing lines of relatively high ionization potential. Late-type stars have spectra showing easily excited lines of low ionization potential and even absorption by molecules such as TiO. Spectral class can be accurately assigned by the line ratios of the different elements in an absorption spectrum.

Each range of temperatures requires specific lines for their accurate determination. For instance, we use lines from $FeII$, HI, $CaII$, FeI, CaI for temperatures in the range $7500 - 4000$ K (spectral types F0 - K5). Lines from $HeII$, $SiIV$, $SiIII$, SII, $FeII$, $MgII$,

Spectral Type	T (K)	Spectral Type	T (K)
O5	45000	G0	6000
B0	30000	G5	5500
B5	15500	K0	4900
A0	10000	K5	4100
A5	8500	M0	3500
F0	7500	M5	2800
F5	6600		

Table 7.1 Relation between spectral class and stellar surface temperature.

	Degree of Ionization							
Atom	I	II	III	IV	V	VI	VII	$VIII$
H	13.6							
He	24.6	54.4						
C	11.3	24.4	47.9	64.5	392.1	490.0		
N	14.5	29.6	47.5	77.5	97.9	552.1	667.0	
O	13.6	35.2	54.9	77.4	113.9	138.1	739.3	837.4
Ne	21.6	41.0	63.5	97.1	126.2	157.9	207.3	239.1
Na	5.1	47.3	71.6	98.9	138.4	172.2	208.5	264.2
Mg	7.6	15.0	80.1	109.3	141.3	186.5	225.0	265.9
Si	8.2	16.3	33.5	45.1	166.8	205.1	246.5	303.2
Ca	6.1	11.9	50.9	67.2	84.5	108.8	127.2	147.2
Fe	7.9	16.2	30.7	54.8	75.5	100	128.3	235.0

Table 7.2 Ionization potentials I_p of various abundant elements (energy in eV; $1\ \mathrm{eV} = 1.602\,10^{-19}$ J). The wavelengths corresponding to ionization potentials follow from $E = h\,c/\lambda$: $\lambda(\text{Å}) = 12413.6/I_p$ (eV)

HI are used in the temperature range $50000 - 7500$ K (spectral types O5 - F0). This method is so sensitive, that one hardly needs to *measure* the line intensity; the presence and appearance of important lines is sufficient.

7.2 The Hertzsprung-Russell Diagram

By plotting for a large number of stars the luminosity L_* as a function of $T_{\rm eff}$, we obtain a diagram known as the *Hertzsprung-Russell diagram*, HRD (see Figures 7.2, 7.4 and 7.3). The great majority of stars occur in a band called the Main Sequence which passes diagonally through the diagram (from hot and bright to cool and faint). There also exist very bright stars over a range of temperatures (top of the HRD: super-giants), bright yet cool stars (to the right in the HRD: red giants) and hot yet very faint stars (bottom left in the HRD: white dwarfs).

Instead of luminosities, we can also plot apparent brightness (magnitudes) in case the stars plotted are all part of the same group, and thus at the *same* distance. Instead of $T_{\rm eff}$, spectral type or colour is often plotted as a function of 'absolute' magnitude $M_{\rm abs}$ instead of luminosity L_*. Magnitude is a traditional astronomical concept:

a. *absolute magnitude:* $M = M_{\rm abs} = \mathrm{cst} - 2.5 \log L$ is a measure for the (distance-independent) luminosity of a star. In the visual (5500 Å), the Sun has $M_{\rm abs} = +4.83$ rendering it inconspicuous among the stars.

b. *apparent magnitude:* $m = \mathrm{cst} - 2.5 \log F = \mathrm{cst} - 2.5 \log L/R^2$ is a measure for the (distance-dependent) flux received from the star on Earth. The Sun's apparent visual magnitude $m_V = -26.74$, makes it the brightest star in the sky.

An unbiased distribution of stars in the HRD requires inclusion of all stars in a certain volume. This is feasible only in the vicinity of the Sun. At greater distances faint stars are overlooked and bright stars become overrepresented.

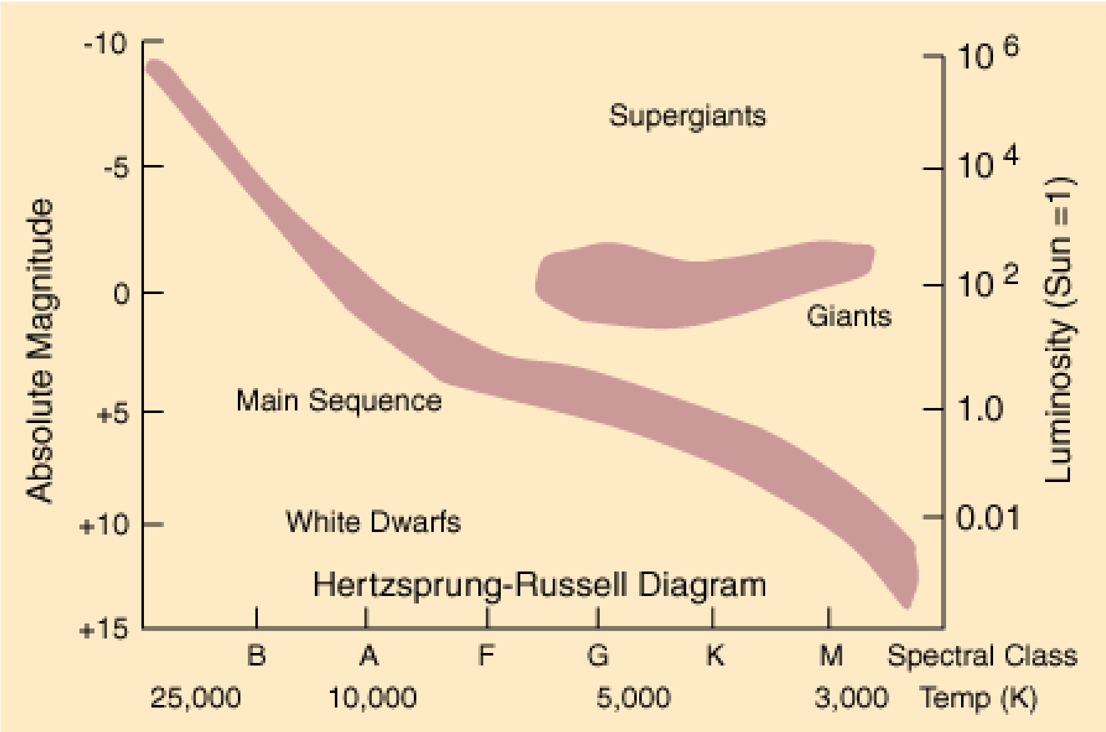

Figure 7.2 Schematic representation of the Hertzsprung-Russell diagram.

7.3 Sizes and Masses of Stars

Consider stellar luminosities and the Stefan-Boltzmann law:

$$L_* = 4\,\pi\,R_*^2\,F_* = 4\,\pi\,r_*^2\,\sigma\,T_{\rm eff}^4\,,$$

with R_* as the distance to the star, r_* the radius of the star and F_* the power received on Earth. Expressing this in Solar units (suffix $\odot$), we obtain a simpler form:

$$\log\left(L/L_\odot\right) = 4\,\log\left(T/T_\odot\right) + 2\,\log\left(r/r_\odot\right)\ .$$

Thus, a relation between L and T automatically implies also a relation between r and T! In this way, the size of a star can be estimated even though in a telescope the star may be undistinguishable from a point source!

Example *The nearby bright star Sirius A is located at a distance of* 8.1×10^{16} m *(*2.6 parsec*; trigonometrically determined). On Earth, we measure a flux of* 1.33×10^{-7} W m^{-2}*. Sirius' luminosity must therefore be:* $L_{\rm Sirius} = 1.1 \times 10^{28}$ W*. Comparing this to the luminosity of the Sun,* $L_{\rm Sun} = 3.9 \times 10^{26}$ W *found earlier, we conclude that Sirius is* 28.2 *times more luminous than the Sun. Its spectral type A0 implies a temperature of* 10000 K*. Thus, Sirius has a radius* $r_{\rm Sirius} = 1.8 r_\odot$.

From the HRD we may determine:

Super-Giants:	$r = 100 - 1000\ r_\odot$
Giants:	$r = 10 - 200\ r_\odot$
Main Sequence:	$r = 0.1 - 10\ r_\odot$
White Darfs:	$r = 0.001 - 0.1\ r_\odot$

N.B.: White dwarfs are no larger than planets in the Solar System, but they are much more massive.

Of all fundamental properties of a star, mass is the hardest to determine. It can be measured by its gravitational attraction, but only when objects move in orbits around

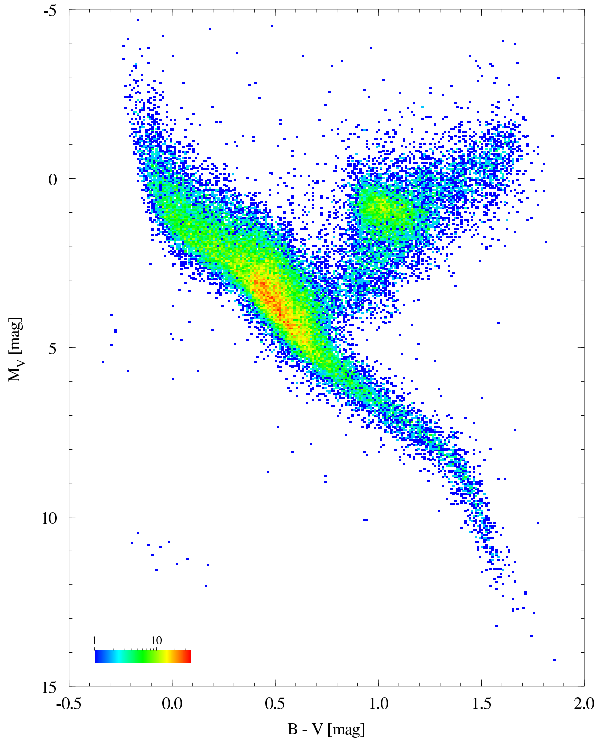

Figure 7.3 Hertzsprung-Russell diagram of 41704 stars precisely measured with the ESA Hipparcos satellite. Instead of log L, absolute magnitudes M_V and instead of temperatures, colours $B - V$ are plotted (blue at left, red at right) From the ESA Hipparcos website

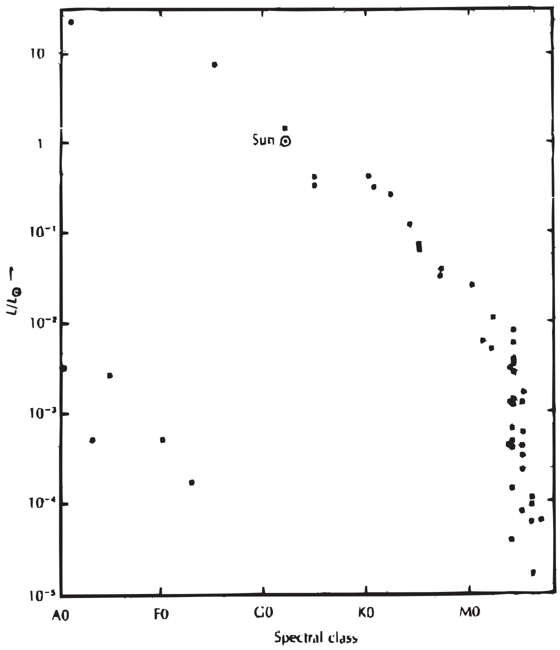

Figure 7.4 The Hertzsprung-Russell diagram for stars within 16 light-years ($5\ \mathrm{parsec}$) from the Sun. Only the two stars Sirius and Procyon are much more luminous than the Sun. Two other stars, α Centauri and τ Ceti are about as luminous, all other stars are fainter, often very much so.

each other (such as binary stars). Usually, many years of measuring binary star positions are required for the determination of accurate orbits from which masses can be deduced with the aid of Newton's law :

$$G\,(m_1\,m_2)\,/\,(r_1+r_2)^2 = m_1\,r_1\,(2\,\pi/\,T)^2\ = m_2\,r_2\,(2\,\pi/\,T)^2\,. \tag{7.1}$$

so that:

$$m_1 = (4\,\pi^2/G\ )[(r_1+r_2)^2\,/T^2]\,r_2\,,$$

and:

$$m_2 = (4\,\pi^2/G\ )[(r_1+r_2)^2\,/T^2]\,r_1\,.$$

and:

$$m_1+m_2 = (4\,\pi^2/G)\ [(r_1+r_2)^3\,/T^2] \qquad (Kepler III). \tag{7.2}$$

From the period T and the two angular distances r_1 and r_2, individual masses m_1 and m_2 may be determined. The distance to the binary must be known in order to convert angular into absolute distances. If only the period T and maximum separation $r_1 + r_2$ (= semi-major axis a) are known, the total mass of the system can be determined, but not the relative contributions of the individual components. Only for nearby binary stars can the distance be determined accurately by triangulation. For more distant stars, approximate distances can be obtained by determining their spectral type and comparing their absolute brightness as predicted by the HRD to the observed apparent brightness.

Tight or distant binaries have angular separations too small to be resolved. Yet, their orbital movement causes a periodic dopler shifts of lines in their (combined) spectrum (spectroscopic binary). The orbital period T equals the period of line-shift and the orbital velocities are deduced from the amplitude or maximum dopler shift. Usually, the inclination i of the orbital plane is unknown, and we can only determine the radial velocity, i.e. the velocity component in the line-of-sight: $v_r\ =\ v\,sin\,i$. Two objects orbiting their common center of gravity *always* have:

$m_1\,r_1\ =\ m_2\,r_2, T_1 = T_2$ and therefore also $m_1\,v_1\ =\ m_2\,v_2$.

From $r = r_1 + r_2 = (T/2\pi)\ (v_1 + v_2)$, and Kepler III, we find:

$$m_1+m_2 = (T/2\pi\,G)\ (v_1+v_2)^3\ = (T/2\pi\,G)\ (v_{r1}+v_{r2})^3\,/sin^3\,i\,.$$

which we transform to:

$$m_1+m_2 = (T/2\pi\,G)\,(v_{1r}/sin\,)^3\ (1+m_1/m_{r2})^3\ ,$$

so that finally:

$$(m_2\,sin\,i^3)/\,(m_1+m_2)^2 = (T/2\pi\,G)\,v_{r1}^3\,. \tag{7.3}$$

The right-hand side of the equation contains observable parameters only, so that the parameters on the left can be resolved. In large samples, we statistically eliminate the $sin^3\,i$ factor by inserting the mean value for randomly oriented orbits.
The actual number of reliable mass determinations is, thanks to the ESA Hipparcos Mission, of the order of a few thousand. As mass turns out to be the most fundamental stellar property, this is still unsatisfactory. Fortunately, diagrams of known (binary star) masses plotted as a function of luminosity reveal an approximate relation (see Fig. 7.7) called the mass-luminosity law:

$$\log\,(L/L_\odot) \approx 3.45\log\,(M/M_\odot) \tag{7.4}$$

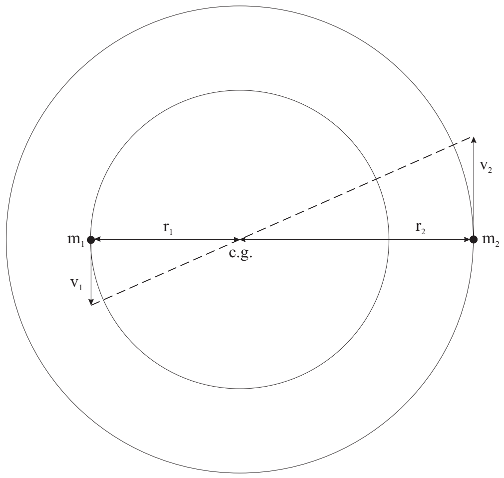

Figure 7.5 The components of a binary star both orbit their common centre of gravity Z.

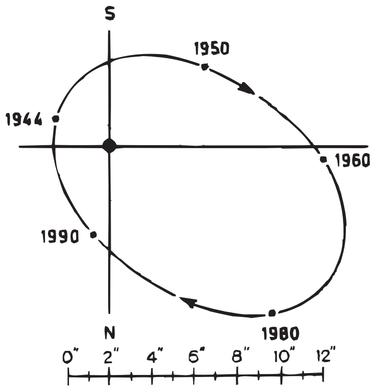

Figure 7.6 Orbit of the binary star Sirius from observations between 1944 and 1990. The scales are in angular measure.

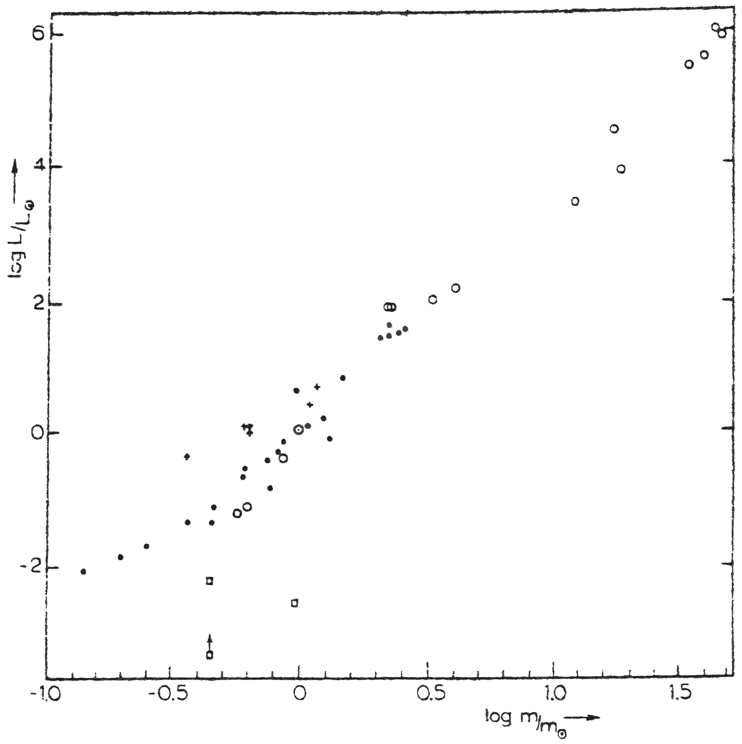

Figure 7.7 The empirical mass-luminosity relation following Eddington. Symbols: o visual binaries; 0 spectroscopic binaries (dopler effect); $+$ binaries in the Hyades open cluster.

Chapter 8

Structure of Sun and Stars

8.1 Origin of Solar Energy

From where does the Sun derive its energy? Some past thoughts:

- The Sun was created hot and has been cooling since. No energy is produced. This can be ruled out. The specific heat of hydrogen is about 5×10^3 J kg^{-1} K^{-1}. A Sun consisting entirely of hydrogen presently at 6000 K has an energy content of only $5 \times 10^3 \times 2 \times 10^{30} \times 6000 = 6 \times 10^{37}$ J. As the Sun loses energy at a rate of (luminosity) 4×10^{26} J s^{-1}, it would have only another 1.5×10^{11} s $= 5000$ years of energy use left. We certainly would have noticed a fairly rapid decrease in the Earth's average temperature over the past centuries.
- The Sun is contracting, and thereby converts potential energy into kinetic energy which degrades to heat allowing energy losses by thermal radiation. This is a more interesting proposition. The amount of energy released by a small mass dm in free fall under the influence of the gravitational field of a total mass M from infinity to a radius r is given by:

$$E_{\mathrm{grav}} = \int_r^{\infty} (G\,M\,dm)\,/r'^2\,dr' = (G\,M\,dm)\,/r\,. \tag{8.1}$$

All masses dm falling freely into a body the size of the Sun release:

$$E = \int_r^{\infty} \left[G\,M\,(r')\;dm\,(r')\right]/r'\,.$$

Assume spherical contraction of the cloud:

$M\,(r) = 4/3\;\pi\,r^3\,\rho$ and: $dm = 4\,\pi\,r^2\,\rho\,dr$. (spherical shell)

Substitute this into the integral, and find for a cloud contracting to radius $r_\odot$:

$$E = \int_{r_\odot}^{\infty} G\,(16/3)\,\pi^2\,\rho^2\,r^4\,dr = (16/15)\,G\,\pi^2\,\rho^2\,r_\odot^5\,.$$

Transform $r_\odot$ back into $M_\odot$:

$$E = (16/15)\,G\,\pi^2\,\rho^2\,r_z^5 = 0.6\,G\,M_\odot^2/r_\odot \tag{8.2}$$

For *inhomogeneous* spherical collapse the result is a factor of 2.5 larger: $E = 1.5\,G\,M_\odot^2/r_\odot$. By substituting actual values, we obtain an energy content $E = 5.5 \times 10^{41}$ J. With a luminosity of 4×10^{26} W the Sun would then be 1.4×10^{15} s $= 4.5$ million years old. Although the physicists Helmholtz and Kelvin were content with this value, it does contradict the 1000 times greater age of the Earth as estimated by geologists.

N.B.: Although the current energy emission of the Sun is not fed by gravitational contraction, this process does play an important role in the creation and evolution of stars: cf. Section 9.2

8.2 Nuclear Fusion in the Solar Interior

The discovery of nuclear fusion in the early 20^{th} century provided the final answer. Einstein's equation $E = m\,c^2$ states that a small amount of matter corresponds to a very large amount of energy. The age of the Sun is about 5 billion years $= 1.6 \times 10^{17}$ s. With its current luminosity, the Sun must have radiated away until now an energy of 6.2×10^{43} J, corresponding to a mass of 7×10^{26} kg, i.e. only $1/3000$ of its current mass.

A star like the Sun converts mass into energy by nuclear fusion, mainly by converting ('burning') hydrogen into helium in its core. At temperatures below 18 million K this occurs mostly in fairly direct conversions (proton-proton cycles). At temperatures above 18 million K it occurs mostly by the carbon cycle which is not to be confused with carbon burning because in this cycle, carbon only serves as a catalyst and not as a fuel. In all these processes 4 hydrogen nuclei (4×1 proton) are converted into a single helium nucleus (2 protons + 2 neutrons). Also released are neutrino's that have such negligible collision cross-sections that they escape from the Sun's interior practically unhindered.

The burning of hydrogen occurs faster at higher temperatures. It is thus fastest in the center, where consequently the conversion of H into He is first completed. Burning of H into He then occurs in a spherical shell around a slowly growing inert He core. With time, the spherical shell moves slowly away from the centre.

For every new helium nucleus, the Sun loses 4.7×10^{-29} kg of mass. Given the mass of a proton ($m_p = 1.67 \times 10^{-27}$ kg), this implies an efficiency of only 0.7%. The mass lost is turned into an amount of energy $E = m\,c^2 \approx 26$ MeV $= 4.3 \times 10^{-12}$ W. Support of the Solar luminosity $L = 4 \times 10^{26}$ W therefore requires almost 10^{38} fusion reactions per second. If the full Solar mass ($M = 2 \times 10^{30}$ kg) were available for hydrogen burning, the Sun's lifespan would be 10^{11} years. In reality, about thirty per cent of the Sun's mass already consists of helium, and much of the Sun is at temperatures too low to support hydrogen fusion (now or at any time in the future). The real lifespan of the Sun is therefore much shorter, of the order of 1.25×10^{10} years of which 40% has already been used up.

N.B. More detail on fusion processes in the Sun is provided in Appendix D

8.3 The Solar Interior

The temperature of the Solar surface was found to be $\approx$ 6000 K. Simple physical considerations predict an increase in temperature and pressure with increasing depth into the Sun. Spectroscopy shows this to be true for the outer layers of the Sun. However, the occurrence of nuclear fusion in the Sun's interior requires very high temperatures and pressures. The evolution of a star like the Sun is governed by the rate at which H is converted into He. Even over the total history of mankind, the Sun has converted only a minute fraction of its mass into energy. Thus, the evolution of the Sun must proceed so slowly that, from a human point of view, we may assume the Sun to be invariable.

This makes it possible to analyze the Solar interior assuming everything to be in equilibrium. It turns out that a valid model requires knowledge of the behaviour of six parameters as a function of distance r to the centre: pressure $p(r)$, density $\rho(r)$, mass $m(r)$, luminosity $L(r)$, temperature $T(r)$ and fusion rate $\epsilon(r)$). This requires the *simultaneous* solution of six time-independent differential equations, such as:

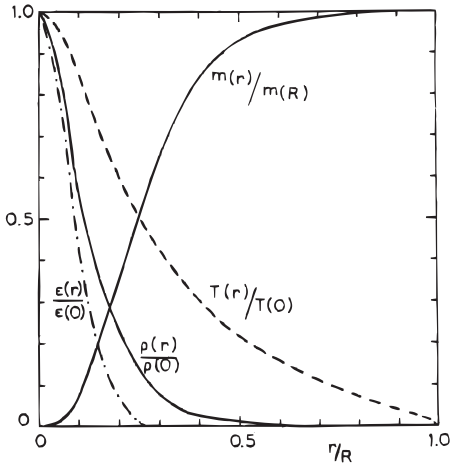

Figure 8.1 Graphical representation of mass m, temperature T, density ρ, and fusion rate ϵ per unit mass as a function of radius r. Suffix R corresponds to the parameter value at the Sun's surface, suffix O to the parameter value in the center of the Sun. From: J. van der Rijst & C. Zwaan, 1975 *Astrofysica*

$$dp/dr = -G\,\rho\,(r)\;m\,(r)\,/r^2 \qquad \text{(hydrostatic equilibrium)} \quad (8.3)$$

$$dm\,(r)\,/dr = 4\pi r^2\,\rho\,(r) \qquad \text{(spherical shell)} \quad (8.4)$$

$$dT\,(r)\,/dr = -\left(L\,(r)\,/4\pi r^2\right)\,(3\kappa)\,/\,\left(16\sigma T^3\right)\;, \quad (8.5)$$

This is valid for radiative transport only. A different equation describes convective transport. Convection never occurs in the hottest stars (O, B), but does in main-sequence stars of type F to M at the transition of stellar atmosphere to burning core. κ is a wavelength-dependent absorption coefficient.

$$dL\,(r)\,/dr = 4\pi r^2\,\rho\,(r)\;\epsilon\,(r) \quad (8.6)$$

In the outer parts of the star, ρ and T are too low for nuclear fusion to occur: $\epsilon\,(r) = 0$ and therefore: $L\,(r) = \mathrm{cst.}$

$$\epsilon\,(r) = f(\rho\,(r\,T(r)\,, composition) \quad (8.7)$$

$$dp\,(r)\,/dT\,(r) = \rho\,(r)\;R/\mu \quad (8.8)$$

R is the universal gas constant. μ is the mean molecular weight; at the very high stellar interior temperatures, all elements are completely ionized (Saha); for complete ionization μ = atomic weight/number of particles: $\mu_H = 1/2$, $\mu_{He} = 4/3$; for all other elements $\mu \approx 2$. Even at very high densities, atomic nuclei and free electrons occupy so little space that the ideal gas law is still valid.

Solving these six equations is not simple, and is usually done numerically on fast computers. Solutions require boundary conditions that specify physical limitations to the mathematical equations. Central boundary conditions are:

$$m\,(0) = 0,\;\; L\,(0) = 0$$

On the other hand, close to the stellar surface ($r \to r_*$), the results need to reproduce actual m and L, as well as:

$$p\,(r_*) = 0,\;\; \rho\,(r_*) = 0$$

8.4 Stellar Characteristics

We may also use the differential equations from the previous Section to evaluate relationships between the various properties of a star. To this end, we assume that all stars have identical *radial* profiles of density, temperature etc., and that they differ only in the *absolute* parameter values.

Start with: $dm(r)/dr = 4\pi r^2 \rho(r)$

which immediately yields: $M \propto \rho_c R^3$

The subscript c marks values corresponding to the stellar centre. M and R are the actual mass and radius of the star.

Substitute $dm(r)/dr = 4\pi r^2 \rho(r)$ into: $dp/dr = -G\rho(r)\, m(r)/r^2$

and integrate over m: $p_c \propto M^2/R^4$,

which we rewrite as: $p_c \propto M^{2/3}/\rho_c^{4/3}$

The ideal-gas equation: $p = n/V\, kT \propto \rho T$

can now be substituted: $\rho_c \propto T_c^3/M^2$

Use $M \propto \rho_c R^3$ to solve for T_c: $T_c \propto M/R$

Thus, already a simple evaluation of the equilibrium equations illustrates several interesting basic properties of a stars. For instance, the larger the radius R of a star with mass M, the lower its central temperature T_c, the lower its central density ρ_c (by a factor of $1/R^3$), and the lower its central pressure p_c (by a factor of $1/R^4$).

We have also seen that for stars on the HRD Main Sequence, specific empirically determined relationships exist between M, R and T_{eff} (*N.B.*: not to be confused with T_c).

From the Stefan-Boltzmann Law: $L \propto T_{eff}^{6.5} \propto T_{eff}^4 R^2$

So that: $R^2 \propto T_{eff}^{2.5}$

From this and the empirical Mass-Luminosity law: $M \propto T_{eff}^{1.9}$

As a result, we find: $R \propto M^{2/3}$

and: $T_{\text{eff}} \propto M^{1/2}$

For the parameter values in the center of a main sequence star we now find:

$T_c \propto T_{\text{eff}}^{3/5} \propto M^{1/3}$
$p_c \propto M^{-2/3}$
$\rho_c \propto M^{-1}$

However, the actual expressions rapidly become much more complicated if we abandon the assumption that all stars have the same overall structure.

Chapter 9

Stellar Evolution

9.1 Evolutionary Tracks

All stellar parameters can be reduced primarily to functions of the *total mass* M_* and the actual *chemical composition* as a function of r. *Only these two parameters* determine all exterior characteristics distinguishing stars, such as T_{eff}, colour, brightness and location in the HRD. However, chemical composition changes over time as a result of internal nuclear burning and stars may lose mass e.g by stellar winds. Assuming semi-equilibrium, changes can be computed in small steps. The results of each computation provide the initial condition for the next, allowing one to numerically follow the whole of stellar evolution. The results of such calculations plotted into a HRD is called an *evolutionary track*.

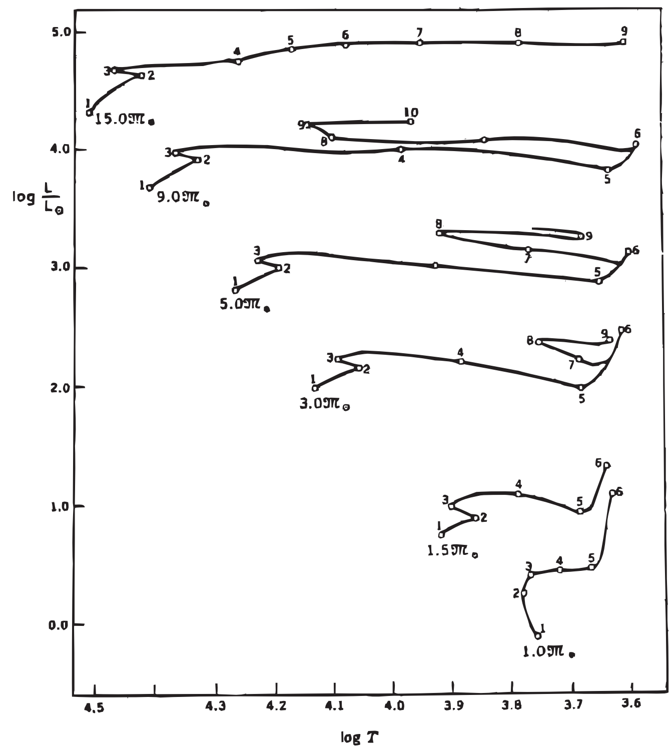

Figure 9.1 Evolutionary tracks in a HRD for stars with identical initial chemical compositions but different initial mass. From: J. van der Rijst & C. Zwaan, 1975 *Astrofysica*

Evolutionary Phase	Path	Mass ($M/M_\odot$				
		15	5	3	1.5	1.0
Contraction to Main Sequence	0 - 1	0.06	0.6	3	20	50
On Main Sequence	1 - 3	10	68	240	2000	11000
Transition to Red Giant	3 - 6	1.5	2	9	280	680
Red Giant	6 - 9	2	20	80	—	—

Table 9.1 Evolution timescales in million years for the stars in Figure 9.1.

9.2 Star Formation

In a tenuous interstellar gas cloud, condensations may form. When such condensations cross the so-called Jeans limit, internal pressures no longer compensate for the gas cloud particles' self-gravity. The cloud will keep on contracting, releasing potential energy (cf. Section 8.1). Gravity increases together with temperature and pressure, and the process never stops. If the collapsing cloud is of mass $M \geq 0.07M_\odot$, central temperatures and pressures eventually become so high that nuclear fusion starts spontaneously. The extra energy released by fusion compensates radiation losses at the surface, and provides sufficient internal pressure to halt the collapse. The object has reached equilibrium and has become a Main Sequence star for millions or even billions of years, depending on its mass. However, further gravitational catastrophe is averted only as long as nuclear fusion keeps on replacing internal energy lost by radiation.

A cloud will contract if the total kinetic energy $E_{\rm kin}$) of all particles in the cloud no longer balances the potential energy U of all those particles relative to one another. From the virial theorem: $2\,E_{\rm kin} < |U|$

In a gas, we have:

$$E_{\rm kin} \;=\; 3/2\,N\,k\,T$$

where the number of particles N is:

$$N = M\,/\,(\mu\,m_{\rm H})$$

M is the cloud mass, μ is the mean atomic weight of the gas particles and $m_{\rm H}$ is the mass of a hydrogen atom. Thus:

$$E_{\rm kin} = 3/2\;M\,k\,T\,/\,\mu\,m_H$$

and the requirement for cloud collapse becomes:

$$3\;M\,k\,T\,/\,\mu\,m_H \;<\; 3/5\;G\,M^2/R \tag{9.1}$$

For a sphere, this is turbed into:

$$k\,T\,/\,\mu\,m_H \;<\; 4\,\pi\,/\,15\;G\,\rho\,R^2$$

We may rewrite this expression as to a critical length for contraction:

$$R > 15\,k\,/\,4\pi\,G\,m_H\,\mu\;\times\;(T\,/\,\rho)^{1/2} \qquad \text{(Jeans Length)} \tag{9.2}$$

or instead a critical mass:

$$M \;>\; (5\,k\,/G\,m_H\,\mu)^{3/2}\,(3/4\,\pi)^{1/2}\;\times\;(T^3/\,\rho)^{1/2} \qquad \text{(Jeans Mass)} \tag{9.3}$$

As we have seen before in Section 8.1, non-homologous free fall collapse releases potential energy to the amount of

$$U = \int_R^\infty E_{grav} = \int_R^\infty G\,M(r)\,dm(r)/r = -1.5\,G\,M^2/R$$

We may use this to roughly estimate central temperatures caused by the collapse. Half of the energy released is turned into heat, the other half is radiated away. A cloud collapsing into a star like the Sun ($M = 2 \times 10^{30}$ kg, $R = 7 \times 10^8$ m) will release $U = 2.75 \times 10^{41}$ Watt sec. If the specific heat of hydrogen is given by $\sigma_H = 5000$ W sec kg^{-1} K^{-1}, we may expect a temperature $T = U/(M\,\sigma_H) = 27.5 \times 10^6$ K. Nuclear fusion will spontaneously occur at such temperatures (and pressures of the order of 10^{11} bar).

9.3 Main-Sequence Lifetimes

We found $\log(L/L_\odot) = \mathrm{cst} + 3.45 \log(M/M_\odot)$. The fuel supply of a star is directly related to M and the fuel consumption to L, so that the life span of a star on the Main Sequence should be proportional to $M/L = \mathrm{cst}\, M^{-2.45}$. Thus, stars more massive than the Sun will have much shorter lifetimes. The most massive stars ($M = 50 M_\odot$) exist for no longer than a few million years! Actual life-spans are summarized in Table 9.2.

Spectral type	Mass ($M_\odot$)	Luminosity ($L_\odot$)	Life-span (10^6 year)
O7.5	25	80000	6
B0	16	10000	12
B5	6	600	65
A0	3	60	250
F0	1.5	6	2200
G0	1	1	11000
K0	0.8	0.4	15000
M0	0.5	0.04	30000

Table 9.2 Mass, luminosity and life-span of stars as a function of their spectral type.

Note that *all stars of types O and B* must astronomically be very young. If one searches for regions of present-day star formation, it is therefore a good idea to look in the vicinity of such stars (for instance, in the constellation of Orion).

In Fig. 9.1 the slow evolution of a hydrogen-burning star on the Main Sequence is shown as track $1-3$. At point 2, all hydrogen in the core is used up and fusion stops. As energy is still radiated away, gravity again overpowers internal pressure and the helium core contracts. This causes internal temperatures to rise so that hydrogen fusion $4\,\mathrm{H} \to \mathrm{He}$ occurs in a shell around the hot, but quiescent core. As a result, the stellar *mantle expands and cools*, increasing the size of the star while the *the core shrinks and heats up* (track $3-6$). Most of this time, stellar luminosities are constant: decreases in T_{eff} are compensated by increases in radius r_*, causing $L = 4\pi r_*^2\, \sigma\, T_{eff}^4$ to remain the same. At the end of the shell-burning phase, moderate-mass stars ($1 - 4M_\odot$) have swollen to red giants, intermediate-mass stars ($4 - 7M_\odot$) have expanded into red supergiants, and very massive stars ($> 15M_\odot$) have become blue supergiants.

9.4 Moderate-Mass Stars, Red Giants and White Dwarfs

The rate of stellar evolution is strongly mass-dependent. The more mass a star has, the more potential energy it turns into heat and the hotter and denser it becomes. The hotter its interior, the faster its fusion engine runs (cf. Appendix D), and the more rapidly it completes each evolutionary stage. The late evolutionary stages are characterized by ever smaller energy gains, making it ever easier for gravity to get the upper hand and for the star to contract. The energy gain per gram matter strongly declines with atomic number. Carbon fusion is less efficient than helium fusion ($\mathrm{He} \to \mathrm{C}$), which in turn is less efficient than hydrogen fusion ($\mathrm{H} \to \mathrm{He}$). Table 9.1 illustrates that stars spend much less time on evolutionary track $3-6$ than on the main-sequence track $1-3$: once a star has quit the Main Sequence, it does not have much time. Transition from one stage to another are accompanied by catastrophic events: stellar pulsations, expulsion of shells, explosions.

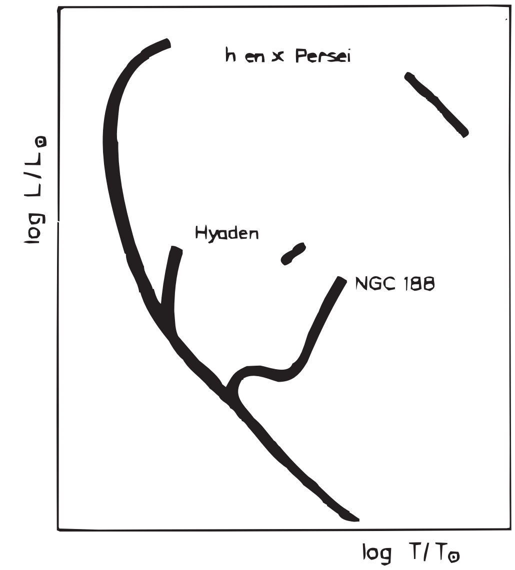

Figure 9.2 Schematic HRD of three open clusters (which is the oldest and which is the youngest?).

The HRD of a group of stars formed at the same time (open clusters, globular clusters) makes it possible to determine the age of the group by finding what stellar masses are still found on the Main Sequence, and what masses occur on the Red Giant Branch. In this way one determines that open clusters (Hyades, Pleiades) in general are fairly young. They have ages of a few hundred million years and even rapidly evolving massive stars are still on the Main Sequence. Globular clusters are very old with ages ≈ 10 billion years. Even slowly evolving stars have had enough time to become Red Giants.

In stars with $M < 4M_\odot$ all processes are relatively slow. The contracting helium core never becomes very hot, allowing it to become extremely dense. It will eventually become 'degenerate' (see Appendix C). As it becomes denser and hotter, the stellar mantle expands and cools (Red Giant phase) until at point 6 on the evolutionary track burning of $\mathrm{He} \rightarrow \mathrm{C}$ starts explosively ('helium flash'). In a matter of minutes about 20% of all helium is converted into carbon. Subsequent energy production is insufficient to prevent further core-shrinking: the star moves along the so-called 'horizontal branch' to a higher temperature at almost constant luminosity. The mass of the contracting carbon core is too small for it to reach the temperatures required for its fusion into heavier elements. When the helium fusion ($\mathrm{He} \rightarrow \mathrm{C}$) draws to an end, the star has turned into a rather small and hot blue-white star, which may blow away a small part of its mass in expanding shells ('planetary nebula'). The star contracts into a white dwarf, made up of degenerate carbon and helium. Its density lies between 10^5 and 10^8 $\mathrm{g\ cm^{-3}}$. Because of its small surface area it cools very slowly, only gradually becoming more red and less bright, and ever more difficult to observe. As it is supported by nonthermal electron pressure (see Appendix C), it does not contract further.

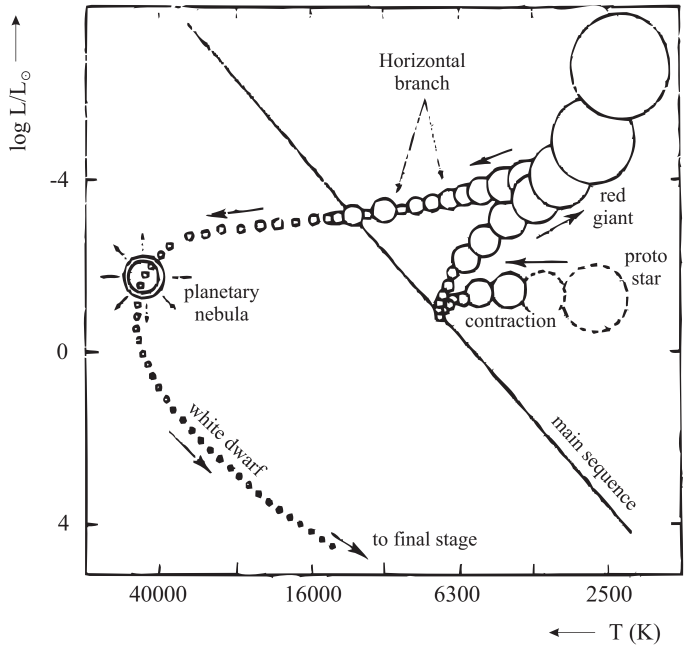

Figure 9.3 Schematic evolutionary track of a star with a mass slightly larger than that of the Sun. The dimensions of the star are not to scale. Schets van De Jager & Van den Heuvel.

9.5 Intermediate-Mass Stars: Red Supergiant and White Dwarfs

Intermediate-mass and massive stars ($M > 4M_\odot$) produce more massive helium cores. Contracting, these easily reach the temperatures (100×10^6 K) required for burning helium: $3\,^4\mathrm{He} \rightarrow\ ^{12}\mathrm{C}$. However, in the core helium is relatively quickly consumed. The inert carbon core contracts and temperatures rise, producing g a helium-burning shell around it. This helium-burning shell is surrounded by a hydrogen-burning shell which in turn is surrounded by a hot, but inert hydrogen mantle.

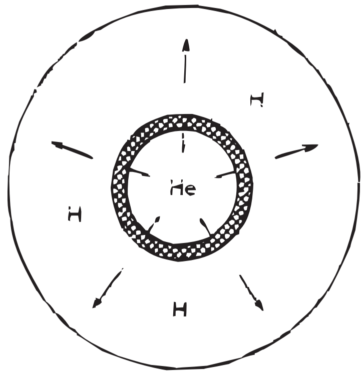

Figure 9.4 Schematic cross-section of a shell-burning star. The inert (hot, but cooling) helium core contracts. In the cross-hatched shell, hydrogen fusion ($4\,^1\mathrm{H} \rightarrow\ ^4\mathrm{He}$) occurs. The hot, but inert hydrogen mantle of the star expands. This and the next Figures are from J. van der Rijst & C. Zwaan, 1975 *Astrofysica*

Eventually, the contracting carbon core becomes so hot that it also 'ignites', converting carbon into yet heavier elements. For stars of mass $4M_\odot < M < 7M_\odot$, this process stops when the core contains a mixture of oxygen, neon and magnesium ($\mathrm{O} - \mathrm{Ne} - \mathrm{Mg}$). After an episode of strong mass-loss caused by violent stellar winds, the star also becomes a

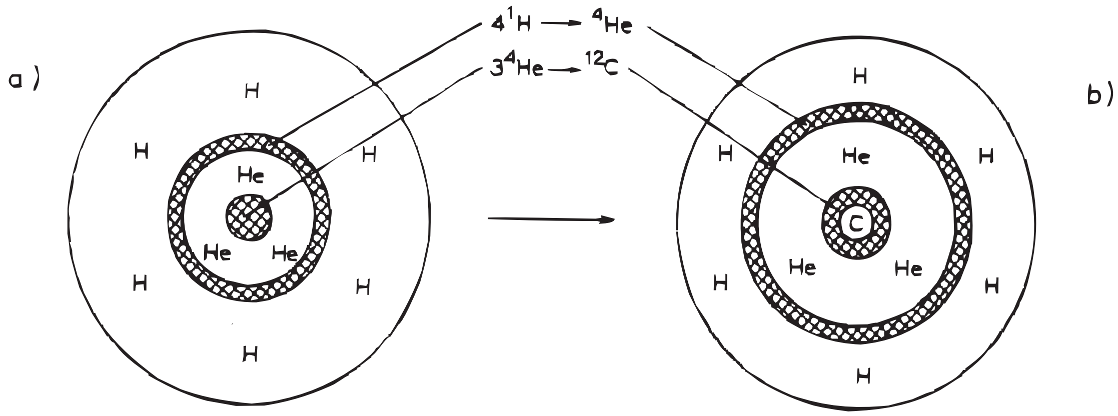

Figure 9.5 Schematic cross-sections of a heavy star ($M > 4M_\odot$) with helium fusion; fusion occurs in the cross-hatched areas. (a) helium fusion in the core, surrounded by an inert helium mantle, in turn surrounded by a shell with hydrogen fusion and an inert hydrogen mantle; (b) at a later stage the star has an inert carbon core, surrounded by a helium-fusion shell, an inert helium shell, a hydrogen-fusion shell and an inert hydrogen shell respectively.

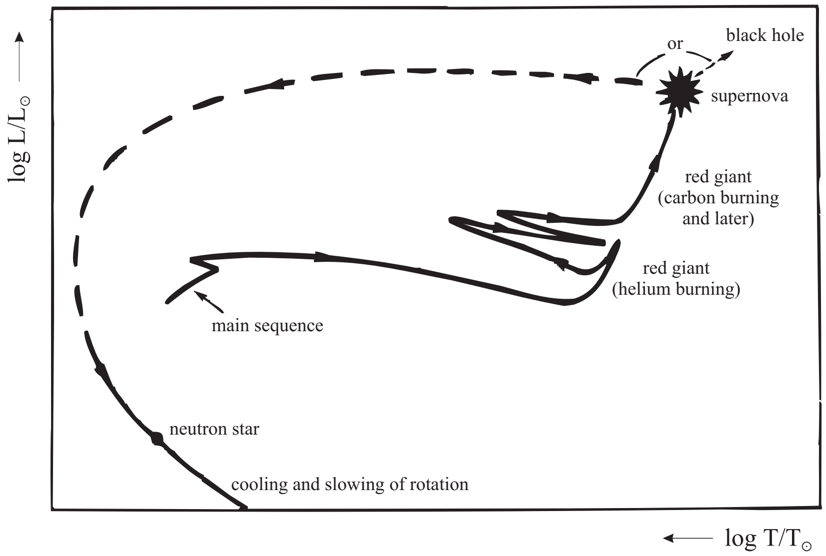

Figure 9.6 Schematic evolutionary track of a heavy star ($M > 4M_\odot$), in this case a star of approx. $7M_\odot$. Schets van De Jager & Van den Heuvel.

white dwarf with a mass of approx. $1.1M_\odot$. If the mass of the stellar remnant exceeds $1.2 - 2.0M_\odot$ (known as the Chandrasekhar-limit), the white dwarf will eventually collapse into a *neutron star* (see Appendix C).

9.6 Massive Stars: Supernova and Neutron Star

In massive stars ($M > 7M_\odot$) nuclear fusion processes continue to produce ever heavier elements until the core consists almost entirely of iron. As iron is at the peak of the binding-energy curve, nuclear fusion will no longer serve as a source of energy. The iron core contracts rapidly increasing temperatures to 5×10^9 K. At such temperatures, the thermal radiation from the core peaks at gamma-ray wavelengths (Section 3.3. The gamma-ray photons are so energetic that they penetrate the Fe nuclei causing them to disintegrate into 13 helium cores and 4 neutrons each. Neutrino's are also produced and carry much of the internal energy of the star away into space. The internal pressure practically drops to zero and the core implodes. From the center of the star, where the imploding matter comes to a halt, powerful shockwaves travel outwards. These shockwaves and the minute fraction of neutrinos absorbed by the mantle turn the implosion into a catastrophic explosion.

This explosive release of energy increases temperatures enormously. For a short period, violent nuclear fusion occurs in the mantle, creating elements heavier than iron. For a short while, high temperatures are also maintained by the radio-active decay of unstable isotopes thus formed. Nearly all of the star, except the core, is hurled into interstellar space at velocities of $10000\ \mathrm{km\ s^{-1}}$: a *supernova* is born.

A supernova is thus an already luminous ($L > 10^5 L_\odot$) supergiant star that for a brief period (weeks/months) increases its brightness even further by factors of approx. 100 *million*. The expanding mantle has a high $T_{\rm eff}$, and a very rapidly increasing surface area. Soon, however, no more energy is produced and $T_{\rm eff}$ drops. The expanding shell grows more tenuous, becomes optically thin, and the supernova luminosity decays. A supernova event returns most of the stellar mass to the interstellar medium, where it eventually may take part in the formation of a new generation of stars (and planets!). The expelled material contains a large fraction of the fusion products, as well as very heavy elements formed in the explosion. Thus, the gas between the stars is enriched in all elements heavier than hydrogen and helium.

Very massive stars are few in number, so that only a small fraction of all stars become supernova. Thus, its appearance is a relatively rare event. In the entire Milky Way, supernovae explode only 2 – 4 times per century. Because the Milky Way is very large and dusty, most of these escape our attention. From Earth, we expect to see such an event no more than once every few centuries. Known stars presently suspected of going supernova 'soon' (tomorrow? in 10000 years?) are the bright stars Betelgeuse (α Orionis), Antares (α Scorpii) and the southern star η Carinae.

The remaining supernova core consists of helium, neutrons and free electrons. It is so dense that electrons combine with protons to neutrons: a *neutron star* is born. The extremely high internal pressure prevents further collapse of the neutron star. Mass densities are of the order of $10^{15}\ \mathrm{g\ cm^{-3}}$. so that a neutron star of mass $1M_\odot$ has a diameter of only several tens of km. Fast-rotating neutron stars (conservation of angular momentum) with mis-aligned magnetic axes produced tightly beamed radio or light pulses (pulsar). As this process also involves energy loss, the rotation of the star gradually slows down.

9.7 Black Holes and Stellar Remains

In a neutron star of mass more than about $2M_\odot$, gravity becomes so strong that not even the neutron-gas pressure can resist further gravitational collapse. This collapse quickly creates a surface gravity so high that escape velocities become equal to the velocity of ligh. Now nothing – *not even photons* – can escape: a 'black hole' has formed, a region of space from where no information can reach the outside world. Inside the black hole, the object that gave rise to it may keep on collapsing – we have no way to tell.

A photon escaping from a gravity well will suffer a gravitational redshift:

$$\lambda/\lambda_0 = [1 - 2\,(G\,M)/(c^2)\,r)]^{-1/2}\,. \tag{9.4}$$

When an object becomes so compact that $[2\,(G\,M)/(c^2\,r)] \to 1$, the wavelength of any photon attempting to escape $\lambda \to \infty$ and its energy $E = h\,c/\lambda \to 0$. This condition defines a critical radius, called *the Schwarzschild radius*:

$$r_S = 2GM/c^2 \tag{9.5}$$

Note that r_S is directly proportional to M. For example, a black hole of mass $10M_{\odot}$, will have $r_S = 30$ km. Because space inside a black hole is distorted, the concept of black hole radius is meaningless. The frequently used *event horizon* is therefore a better designation. However, 'radius' is still useful in defining:

$$2\,\pi\,r_S = 4\,\pi\,G\,M/c^2 \qquad \text{(circumference)}$$

$$4\,\pi\,r_S^2 = 16\,\pi\,G^2\,M^2/c^4 \qquad \text{(surface area)}$$

9.8 Interstellar Gas Enrichment

At the end of their life, stars return part of their mass to interstellar gas and dust clouds from which new stars may form. The exploding supernova ejects large amounts of fusion products. All iron, and elements heavier than iron such as lead, uranium, cobalt etc., have originated in explosions of massive stars. On a milder scale, ejected planetary nebula shells also contain products of stellar fusion, while red giants and supergiants have fierce stellar winds blowing from their atmospheres, and in this way also enrich the interstellar medium with elements lighter than iron. These less massive stars are, in fact, mostly responsible for the relatively abundant elements carbon, nitrogen and oxygen. The Earth, and all life on it, mostly consists of material that has at one time been part of a stellar interior: 'we are star-stuff'.

The existence of a practically universal 'cosmic abundance' of 90% hydrogen and 10% helium with only a small pollution shows that the universe has not been in existence long enough to enable stars other than the – rare – relatively massive stars to return their nuclear fusion products to interstellar space. The current cosmic abundance of H and He is only marginally different from the 'initial composition' of the universe. When we compare the mass-dependent life expectancies of the stars with the current age of the universe (≈ 14 billion years), it is obvious that only stars with initial masses larger than that of the Sun have reached their final phase. All stars less massive than the Sun that were ever formed, are still around.

The final remnants of stellar evolution (white dwarf, neutron star and black hole) are very stable. As they no longer produce energy, they radiate little and are very difficult to detect. If the universe remains in existence long enough, they will eventually 'vaporize'. However, this only happens on very long time scales.

Chapter 10
Other Worlds

10.1 Formation of the Solar System

Out of a contracting and fragmenting interstellar gas cloud, perhaps encouraged by a nearby supernova, the Sun formed 5 billion years ago. The planets followed ≈ 0.5 billion years later. The net angular momentum (rotation) of the contracting cloud caused it to collapse much faster parallel to the spin axis than perpendicular to it; it rapidly turned into a thin disk. The central proto-Sun caused a radial temperature gradient in the 'primordial cloud'. Cloud density and pressure also varied from relatively high in the center to relatively low at the outer edges. Depending on temperature and pressure, gases condensed into solid material ('freeze out') or formed various molecules. Turbulence and viscosity led to the formation of small condensations in the orbiting disk. By low-speed collisions, molecules could accrete into particles and even large solid grains. By repeated collisions and mergers, small condensations grew into ever larger objects. The planets finally grew by accreting many small asteroids. Many objects throughout the Solar System have crater-saturated surfaces, as testimony to the intense bombardment of planets and moons by smaller objects in the first billion years of the Solar System. The formation of the planets has not yet quite come to an end. Accretion still occurs but it has – fortunately – become a rare event as few potential impactors remain. The extinction of dinosaurs and plankton about 65 million years ago was the result of an impact by an asteroid some 10 km across. More recently, in June 1994 the planet Jupiter was hit hard by more than a dozen kilometer-sized pieces from a comet that it had captured and tidally demolished a little earlier.

10.2 Planetary Composition

Temperature-pressure chemical equilibrium models of the primordial cloud tell us which materials formed and solidified at various distances from the Sun and participated in the creation of the protoplanets. For instance, iron became a pure solid around the orbit of Mercury. At the Earth's orbit, it all became bound in iron-sulfide, and near Mars it was oxidized to iron-oxide. Between the orbits of Jupiter and Saturn, water vapour turned into ice. The models are complicated by the fast formation of the planets in only a few hundred million years. Especially in the cold outer parts of the Solar System reaction rates are low and not all chemical reactions could have reached equilibrium. However, more complex models do not give fundamentally different answers.

In fact, the model results are astonishingly good. The Mariner 10 spacecraft found Mercury planet to consist of a huge iron core covered by only a thin moon-like mantle of silicate rocks. Venus and the Earth probably once were very similar. Venus is mostly solid and does not have an internal magnetic field. However, the Earth condensed at a lower temperature so that, unlike Venus, its core is rich in iron-sulfide and remains liquid at relatively low temperatures making possible the long-term existence of a strong magnetic field. Also unlike Venus, the Earth retained water-containing silicates, and it has large quantities of water in the oceans. Mars also has an iron-sulfide core and its red colour is direct proof that it has a silicate rich in iron-oxides. It should also have retained water-rich silicates and there should be large amounts of water on the planet. Current guesses are that it may reside in the soil as permafrost or as large water-ice deposits at the poles. This is a strong driver for the presently intense exploration of Mars. If so much water can be found on Mars, it might even become possible to 'terraform' the planet ...

Between the orbits of Mars and Jupiter, in a wide gap where another planet might be expected, a large number of much smaller objects, asteroids, orbit. Perturbations caused by the nearby and massive Jupiter precluded formation of a full-blown planet. Fewer than 200 asteroids are larger than 100 km; the total mass of all asteroids is less than $1/3000$th of the mass of the Earth. Models predict that the asteroids should mainly consist of water-rich rock, but asteroids made of iron also turn out to exist. Asteroid formation must have been a complex process. In the inner Solar System, planets were unable to retain their volatile elements and they were quickly driven out of the system by photon pressure and particle winds from the Sun. The speed of particles in a gas is proportional to the gas temperature ($v_{\text{avg}} = (8kT/\pi m)^{1/2}$) and it is only at Jupiter's orbit and beyond that gas particle speeds were below (proto)planet escape velocities.

The more distant planets of the outer Solar System thus succeeded in retaining their volatiles. As these planets have masses between a 100 to a 1000 times greater than the Earth, we can easily guess what we are missing ... As the models predict, the outer gas giants have small rocky cores embedded in deep and massive atmospheres consisting of molecular hydrogen, helium, ammonia and methane. Temperature and pressure increase with atmospheric depth, to such extent that molecular hydrogen becomes liquid, or even metallic. Except Saturn's Titan, target of the US-European space mission Cassini-Huygens, the moons of these planets lack the mass to retain an atmosphere even at the low ambient temperatures. Ices are a major constituent. Inside some moons (e.g. Jupiter's Io and Europa) tidal friction produced by the planet and neighbouring moons causes significant internal heating. The Voyager spacecraft flyby's showed several active sulfide-geysers on Io, and a sludgy ocean underneath a continuously cracking and refreezing icy surface on Europa.

10.3 Planets Orbiting Other Stars

The above suggests that planets orbiting another star should be small and rocky close to the star and gaseous and massive far from it. Recent results have considerably upset such expectations. Other stars do have planets. Observations with infrared satellites (IRAS, ISO, SIRTF) already revealed many stars with near-IR (NIR) and mid-IR (MIR) emission in excess of what the stars produce themselves. This excess is ascribed to dusty disks radiating at planetary temperatures. HST images of e.g. the Orion nebula show heavily obscured stars surrounded by opaque disks seen in silhouet against the nebular light. Many of these disks may represent failed or failing planetary system formation, but they show that this process occurs frequently among stars. Actual planets orbiting other stars may be detected in various ways.

- Direct observation is difficult because the reflected light of a planet is much fainter than the light of the much brighter illuminating star. The difference is least at wavelengths close to the emission peak of the planets (i.e. NIR/MIR). No extrasolar planet has yet been seen.
- If the orbital plane of an extrasolar planet lies in the line-of-sight, it will periodically occult its parent star. During the transit, we may detect the resulting small decrease in stellar brightness. Very accurate photometry of such an event has been obtained in the case of an extrasolar planet already found by other means (see below).

- The movements of an invisible planet and a visible star around their common center of mass may be detected by very accurate astrometry as the sinusoidal path of a star on the sky. By this method, small stellar companions have been found, but not (yet) planetary companions.
- The orbital velocity component in the line-of-sight (the *radial velocity* will cause a dopler effect that is seen as a back-and-forth movement of stellar absorption lines. The period of this movement corresponds to the period of revolution T of the planet. The amplitude is proportional to the planetary mass m_2 and inversely proportional to the planet's distance to its parent star r_2.

.

We will look at this last case in more detail (cf. Section 7.3). From Newton/Kepler III (Eqn. 7.6):

$$m_1 + m_2 = (4\,\pi^2/G)\;(r_1 + r_2)^3\,/T^2, . \tag{10.1}$$

Because the planetary mass m_2 is much smaller than the stellar mass m_1, the distance $r_1 << r_2$. Thus, neglecting both m_2 and r_1, we find:

$$r_2 = [(G\;\left(m_1 + T^2\right)/(4\,\pi^2)]^{1/3}\,.$$

Because $T = 2\,\pi\,r_1/v_1 = 2\,\pi\,r_2/v_2$ we have

$$v_2 = [2(\,\pi\,(G\,m_1)\,/T]^{1/3}\,.$$

And with $m_1\,v_1 = m_2\,v_2$ this becomes

$$m_2 = m_1\,v_1[T/(2\,\pi\,(G\,m_1)]^{1/3}\,.$$

From the spectral type, we may estimate the stellar mass m_1 (for instance, Solar-type stars have $m_1 \approx 1\ \mathrm{M}_\odot$). However, instead of the velocity v_1 we have observed only the radial velocity $v_{1,r} = v_1\,sin(i)$, where i is the inclination of the orbital plane to the plane of the sky. We finally obtain:

$$m_2\,sin(i) = m_1\,v_{1,r}[T/(2\,\pi\,(G\,m_1)]^{1/3}\,. \tag{10.2}$$

At the right, we now have only known factors; the product $m_2\,sin(i)$ can evaluated statistically assuming random inclinations. If the inclination can actually be determined ($i = 90^o$ by detecting an occultation; any i by careful astrometry) the actual mass follows.

This method requires very accurate spectral line measurements which have become possible only since 1995. Present accuracy is up to 1 m/sec (i.e. 10^{-8}). Over a thousand stars have been observed, including virtually all Solar-type stars within 30 lightyears from the Sun. So far, more than hundred planets have been found. Most of them turn out to be very massive, comparable to Jupiter, and to orbit very close to their parent star. This is mostly the result of a strong bias inherent in the detection method. As such planets cause the largest spectral line dopler shifts in the shortest periods of time, they are the easiest to discover. Nevertheless, any planetary system with a gas giant in an inner orbit is going to be very different from the Solar System. These gas giants appeared to have formed farther out and to have migrated inwards. Any smaller planets in terrestrial-planet orbits must long ago have been ejected or absorbed by the giant. Many of the planets found also have very elliptical orbits, causing other planet's orbits to be periodically disturbed. Thus, our Solar system appers to be very different from all of the hundred or so other systems found to date. However, only now do we start to have the accuracy and time base to find actual counterparts of our system. Whether the Solar System is exceptional or normal is an open question, which will be answered in the coming decade or so.

10.4 Conditions on Other Worlds

We may deduce at least some of the conditions on planets orbiting other stars (cf. Section 4.2. For a rapidly rotating planet with a decent atmosphere, we found:

$$T^4_{\mathrm{planet}} = 1/4\;T^4_{\mathrm{eff}}\,(r_\star/R_{\mathrm{planet}})^2\,(1-a)\;, \tag{10.3}$$

Considering habitable planets around other stars of different mass and temperature, we require a comfortable surface temperature: $T_{\rm planet} = T_{\rm Earth}$, which translates into an orbital distance $R_{\rm planet}$ obeying:

$$T_*^4\,(r_*/R_{\rm planet})^2\,(1-a_{planet}) = T_\odot^4\,(r_\odot/R_{\rm Earth})^2\,(1-a_{Earth})\ ,$$

from which follows:

$$R_{\rm planet} = (T_*/T_\odot)^2\,(r_*/r_\odot)\ [(1-a_{planet})\,/\,(1-a_{Earth})]^{1/4}\ R_{\rm Earth}\ . \tag{10.4}$$

The orbital period ('year') of the planet will be:

$$T = 2\,\pi\,[R_{\rm planet}^3\,/\,(G\,M_*)]^{1/2}\ . \tag{10.5}$$

The colour and intensity of 'sun'light will also be a function of the star's temperature and distance. By way of example, we imagine an A0 star ($T = 10000$ K, radius $r = 2.5 r_{sun}$) and an M0 star ($T = 3500$ K, radius $r = 0.63 r_{sun}$. A twin of the Earth would orbit the A0 star at 7.4 AU and the M0 star at 0.23 AU. As the masses of these stars are 3.2 and 0.47 M$_\odot$ respectively, the planets will have 'years' of 12.7 year and 0.16 year = 59 days respectively. Using Planck's law to determine the brightness of 'sun'light to our eyes, we find a surprising result. The first planet is relatively faintly and blueishly lit by its A0 sun, with only 0.15 times the brightness of the Sun on Earth. The second planet is reddishly lit by the M0 sun, with 0.7 times the Sun's brightness. Any star unlike the Sun seen from a habitable planet has a brightness less than the Sun seen from Earth. This effect is stronger with increasing spectral (temperature) difference. We would probably notice the colour difference (although we may get used to it easily). However, we may not notice the brightness difference, as the human eye interprets intensities logarithmically. Both stars will still be too bright to look at directly.

How big would the stars be in the sky of these planets? As $r_*^2\,T_{eff}^4 = {\rm cst}$, the A0 star would be a third of the Sun (11′) but the M0 star would be three times bigger, subtending an angle of 1.5^o. Life on the planet orbiting the A0 star would not be pleasant, because this star emits a large fraction of its energy at life-threatening UV-wavelengths. Nothwithstanding its greater distance, the A0 star would irradiate its planet with much stronger UV-radiation than the Sun showers on Earth
Stars brighter than the Sun are surrounded by a relatively broad zone in which habitable planets may occur, but they evolve more quickly. Their more rapid luminosity increase shifts the habitable zone relatively quickly past the orbit of any habitable planet. If planets do not occupy an habitable zone for more than a few billion years, there appears hardly enough time for any serious evolution of life. Stars much fainter than the Sun evolve very slowly, but they have only very narrow habitable zones, again creating an unfavorable net effect. Is it coincidence that the situation is optimal for stars of spectral type F5 to G8, with the Sun as a G2 star just in the middle?

Chapter 11
Interstellar Medium

11.1 Inventory

No more than 6000 stars are visible to the naked eye. However, even the smallest telescope shows many more. In a telescope the faint and diffuse band of light in the sky, the Milky Way, also dissolves into numerous stars. From the darkness of the sky at night we conclude that the Sun belongs to a system of stars that is finite in extent. The gradual discovery of this system, our Galaxy which we call the Milky Way, occurred over the last century. Only a small part of it can be seen directly by us. Almost everything more distant than about 3 kiloparsec = 10000 light-years is visually completely obscured by the high optical depths caused by the presence of both diffusely distributed dust and dense dust clouds. The dust extinguishes not only visible light but also near-IR and UV radiation. The neutral hydrogen with which the dust is mixed also completely absorbs far-UV radiation and soft X-rays. Hard X-rays, gamma-rays, as well as mid-IR, far-IR and radio waves propagate (mostly) unhindered through the Galactic gas and dust. However, very long-wavelength radio waves are scattered completely by Galactic gas.
The Galaxy does not only contain stars, gas and dust. A survey of its contents includes:

- photons from:
- stars and stellar remnants, and:
- luminous emission nebulae, often associated with:
- reflecting and obscuring clouds of dust mixed:
- diffuse neutral gas, both atomic and molecular;
- energetic cosmic ray particles accelerated by:
- magnetic fields;
- dark matter of unknown origin

Most of these items together make up the *interstellar medium*.

11.2 Hydrogen

The Cosmic Abundance (see end of Chapter 6) implies that much of the interstellar gas must be hydrogen and helium. Indeed, dust particles, molecules and other elements only form a minuscule, but often very important, additive to the interstellar medium.

We consider the most common element, hydrogen, in more detail. It occurs in the interstellar medium in various guises:

1. **Neutral, atomic hydrogen (**H **or** HI**)** is the most common form. Its easily observable radio spectral line at a wavelength $\lambda_{HI} = 21.1$ cm ($\nu_{HI} = 1421$ MHz) is a fine-structure line that represents the small energy difference $\Delta E = h\,\nu_{HI}$ between an H

atom in an upper state where proton and electron rotate in the same sense (parallel spin) and a lower state where the rotations are in the opposite sense (anti-parallel spin). Atoms in the upper state 'spontaneously' (mean waiting time several million years!) decay to the lower state emitting a low-energy photon. Neutral hydrogen is only visible at radio wavelengths.

Typical particle density: $0.1 - 1\ \mathrm{cm}^{-3}$.
Typical (spin) temperature: 100 K.

2. **Ionized, atomic hydrogen (H^+ or $\mathrm{H}\,II$)** occurs close to stars of spectral type O, hot enough to emit large numbers of UV photons in the Lyman continuum ($\lambda < 912\,\mathrm{\AA}$). Such UV photons carry sufficient energy to strip an electron from its proton (hydrogen ionization potential: 13.6 eV). Recombinations in ionized hydrogen clouds produce a weak free-free continuum but strong *emission lines* in the Lyman series (UV), the Balmer series (visual) and the Paschen and Brackett series (near-IR). Radio recombination lines (bound-bound transitions at high n and m levels) are also easy to measure. The intense Balmer lines in the visual, and especially the strongest Balmer line $\mathrm{H}\,\alpha$ at 656.3 nm, make for beautiful red colours in nebula images; green colours are usually caused by strong oxygen (O^{++}) emission, whereas blue corresponds to starlight reflected on dust by Rayleigh-scattering.

   Typical particle density: $10 - 10000\ \mathrm{cm}^{-3}$.
   Typical (electron) temperature: 10000 K.

3. **Neutral, molecular hydrogen (H_2)** is formed at high gas densities ($n_\mathrm{H} > 500\ \mathrm{cm}^{-3}$) when H atoms meet on the surfaces of dust particles. It is usually found in large and dense cloud complexes. The Milky Way contains some 4000 of such complexes. Because it is a symetrical molecule, H_2 does not have a dipole moment and is difficult to observe. Quadrupole transitions in the far-UV can only be observed by satellite (NASA: Copernicus), but these UV observation are severely hampered by interstellar extinction. Numerous near-IR emission lines can be observed from mountain observatories. They originate in 'fluorescent' or in shocked H_2 gas heated to $1000 - 2000$ K, which represents unfortunately only a minute fraction of all H_2 gas. Finally, H may occur as part of other molecules – see below.

   Typical particle density: $1000 - 10000\ \mathrm{cm}^{-3}$.
   Typical (kinetic) temperature: $10 - 100$ K.

Next to hydrogen (and helium), elements such as oxygen, nitrogen and carbon are also relatively common in the interstellar medium, but much less than hydrogen.

11.3 Dust

Dust particles are *formed* in the (cool – 1000 K) atmospheres of red giant and supergiant stars. At their formation, dust particles are about $0.1\,\mu\mathrm{m}$ in size. In dense clouds they may *grow* to larger sizes (accretion) by soft collisions. When they encounter shockwaves in the more diffuse interstellar medium, they *shrink* by erosion, and if they come very near hot stars, they *'evaporate'* alltogether at temperatures of $T > 1500$ K. Interstellar dust particles probably consist of loose collections of silicate cores with mantles rich in carbon and water. The cores are more resistant than the 'fluffy' mantles. Accretion and erosion together result in a population of dust particles of various sizes. Dust causes interstellar extinction, which is nil at radio wavelengths and small in the IR but rapidly increases in the visual through the UV region. In the Solar Neighborhood, dust masses are about $0.25 - 1.0\%$ of the associated gas masses.

11.4 Molecules

Molecules may be formed by encounters in a gas, or on the surface of dust particles. In either case, sufficiently high densities ($>> 100\ \mathrm{cm}^{-3}$) are required. Gas-phase molecule formation occurs in complex networks of chemical reactions, which can only be modeled by computer. The most common molecule after H_2 is carbon monoxide (CO). Other often-encountered molecules are generally those constructed from the most abundant elements:
OH (hydroxyl), CN, CS, NO,
H_2O (water), HCN, HCS, SiO, HCO^+,
NH_3 (ammonia), H_2CO (formaldehyde), SO_n, C_nH_m
etc. Even more complex molecules such as alcohols and amino acids have been found. Abundances with respect to H_2 are invariably low, of the order of $< 10^{-6}$ or less! Emission from some molecules such as OH and H_2O can become very intense by stimulated emission (*maser radiation*). Most molecules have spectral lines which can be observed with ground-based observatories at radio cm and mm wavelengths.

Since *cool* molecular hydrogen is not easily observed from the Earth's surface, the existence of molecular clouds went unnoticed for a very long time. Only observations of other less abundant molecules, in particular the 'tracer' molecule CO, have changed this since 1975. In the Milky Way, the CO abundance with respect to H_2 is about 10^{-5}. Many of its emission lines (wavelengths of $2.6/n$ mm; n = 1, 2, 3 ... etc.) are observed with ground-based millimeter telescopes such as the 15 m JCMT, UK-Can-NL telescope in Hawaii). Observations at the shorter wavelengths have to be performed from space (Herschel - ESA cornerstone mission).

Molecular clouds consist of clumps and filaments with sizes varying from 0.3 to 30 light-years and masses varying from 1 to 1000 solar masses. The clouds form large complexes with sizes up to 500 light years and masses up to 10^6 or 10^7 solar masses. Measurements of infrared emission from radiating cool dust in such complexes (e.g. with the US-NL-UK satellite IRAS) have shown that these cloud complexes are the birthplace of most, if not all, stars. One of the nearest such is found in the constellation Orion, where numerous young, hot stars are closely associated with a large molecular cloud complex, in which stars are still being formed at present. The famous Orion nebula is like a 'blister' on one of those molecular clouds. Its caused by a few hot stars at the edge of the cloud, formed only a million years ago. These stars are eroding by ionization their 'birthcloud'.

11.5 Cooling of Gas Clouds

Ionized hydrogen (HII) nebulae, heated by the UV radiation of embedded ionizing stars, invariably have (electron) temperatures T_e between 5000 K and 13000 K. Temperature equilibrium is reached by cooling balancing the heating. The only cooling mechanism available to gaseous nebulae is energy loss by emitting photons via atomic transitions (in particular bound-bound transitions. The relatively common elements C, N and O have many more of such transitions than the very simple hydrogen atom, and are easily excited by collisions with other gas particles. The rate at which collisions occur is, of course, proprtional to the temperature. When an excited electron falls back to its lower level, the kinetic (thermal) energy gained in the collision is emitted as a photon which can escape the gas cloud. As the result is cooling of the cloud, it can easily be seen that elements such as C, N and O in a HII region in fact work as a thermostat!

Emission lines important for cooling the interstellar medium are CI (at 608 and 371 μm in the sub-millimeter region), CII (at 157 μm in the far-IR) and OIII (especially the 88 μm line, likewise in the far-IR). With considerable effort, these lines can be observed from airborne observatories (planes, balloons), but much better from spacecraft (ESA ISO; NASA SIRTF).

11.6 Shockwaves

The interstellar medium is heated, not only by stellar photons but also 'mechanically' by shock-waves. Interstellar shock-waves may result from:

- **Ionization**: when interstellar gas is first ionized, its temperature suddenly rises from ≈ 100 K (neutral) to ≈ 10000 K (ionized), initially without change of volume. As the ideal gas law applies, a large and sudden *increase in pressure* must occur, which causes a shock-wave in the adjoining medium.
- **Stellar Mass Loss**: any loss of mass from stars implies a gas flow away from the star. The outflowing gas collides with the surrounding medium. As the sound speed in cold neutral gas is only a few km/sec, shock-waves usually result.
- **Density Waves**: the motion of stars around the center of a galaxy produces rotating, spiral-shaped density waves, which are preceded by a shockwave.

When a gas cloud encounters a shock-wave, the 'ordered' kinetic energy of the cloud motion relative to the shock is conserved but (i) converted to the 'chaotic' kinetic energy of heat. The strongly heated gas behind the shockwave (ii) rapidly cools by radiation thus reducing internal pressures (ideal gas law). The cooling gas (iii) then collapses to high densities. In doing so, the cloud (iv) may exceed the Jeans citerion (see Section 9.2) in which case unrestrained contraction leading to star may occur. In fact, the formation of our Solar System may have been initiated in just this way by an expanding shock wave caused by a supernova explosion 5 billion years ago.

Chapter 12

The Milky Way Galaxy

12.1 Overall Structure

Because of significant interstellar extinction in the plane of our Galaxy, we cannot see very far optically. Nevertheless:

- the Galaxy must be *finite*, since the sky is dark at night; if this were caused by interstellar extinction only, it would be very hot on Earth.
- the Galaxy must be flat *flat*, because the visible stars are not evenly distributed across the sky. The Milky Way is a band on the sky, and stellar space densities decrease rapidly away from the Galactic plane.
- the Galaxy has a *center* located in the constellation Sagittarius, at a distance of about 28000 light-years (current value). Visually, the center is obscured by dust clouds (visual optical depth $\tau_V = 25$. This was first deduced from the distribution of *globular clusters*. There are over 100 globular clusters, each containing roughly 100000 stars. Their HRD shows that they are very old and their their distance to us is with the aid of RR Lyrae variables, stars pulsating with a period closely related to their luminosity. Their apparent brightness compared with the luminosity deduced from their period reveals their distance. Globular clusters were found to be distributed spherically around the center of the Galaxy, betraying its location.
- most stars are located close to the Galactic plane, but there is a population of very old stars also spherically distributed around the Galactic center: the *halo* of the Galaxy.

In 1944 Van de Hulst predicted that neutral hydrogen atoms H I would emit radiation at a wavelength of 21 cm. This radiation was subsequently detected in the US, Australia and the Netherlands (Kootwijk). With a 25 m telescope in Dwingeloo (completed in 1956) and similar telescopes in the Southern Hemisphere, the 21 cm H I line emission has been mapped over the whole sky.

- H I is found everywhere in the Galaxy;
- 21 cm radio radiation is not absorbed by dark clouds, so that H I can be '*seen*' throughout the Galaxy;
- H I emission is line radiation, allowing not only the measurement of intensity as a function of direction, but also velocity relative to the Sun of the emitting gas anywhere in the Galaxy.

By measuring H I cloud velocities and combining these measurements with a *rotation model* of the Galaxy, we may determine for each cloud its distance to the Sun and its distance to the Galactic Center. From maps so created it appears that our Galaxy, like many other galaxies, has a pronounced spiral structure (see Figure 12.1). A simple version of the rotation model is obtained by determining maximum velocities as a function of longitude θ (angle between viewing direction and the Galactic Center) and assuming circular orbits.

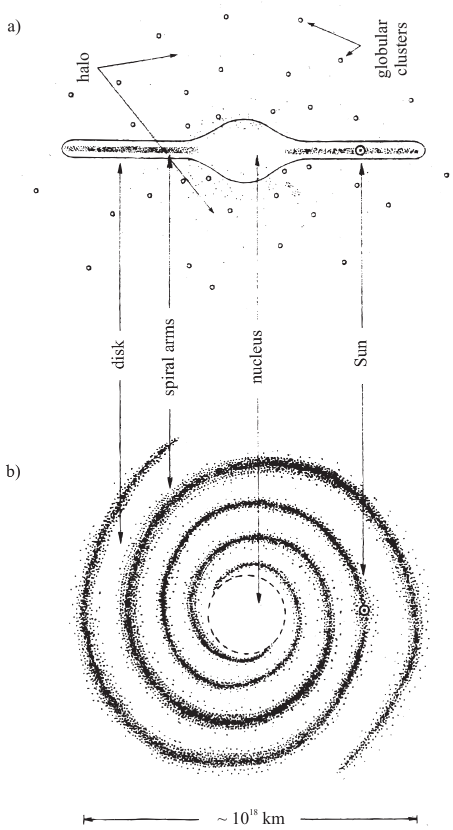

Figure 12.1 Schematic representation of the Galaxy; (a) side view (b) view from above. The Galactic disk with its spiral arms is surrounded by a spherical halo also containing the globular clusters. The location of the Sun is indicated by $\odot$. From J. van der Rijst & C. Zwaan *Astrofysica*.

The line of sight is always tangential to the orbit of the gas associated with this maximum velocity. This velocity $v_{max}(\theta) = v(\theta)$ is the combined motion of gas at a distance $R = R_o \sin\theta$ from the Galactic Center and the Sun moving at distance R_o with velocity $v_o(R_o)$. The circular velocity of the gas in orbit R is $v(R) = v_{max}(\theta) + vo(R_o)\ sin(\theta)$ In this way one obtains the so-called Galactic *rotation curve*, the relation between V and R throughout the system. Rotation curves are even more easily obtained for other galaxies which we observe from the outside, obviating the need to determine the Galactocentric motion and distance of the Sun.

12.2 Rotation Curve

The shape of a rotation curve contains valuable information about the mass distribution of a galaxy. Once again assuming circular orbits:

$$v(R)^2 = G\,M(R)/R\,, \tag{12.1}$$

where $M(R)$ represents all mass *inside* the radius R. The mass outside this radius has no influence on velocity. If we write the radial mass distribution as $M = M_0 R^\alpha$, we find: $v(R)^2 = G\,M_0 R^{\alpha-1}$, or:

$$v(R) = (G\,M_0)^{1/2}\left(R^{(\alpha-1)/2}\right)\,. \tag{12.2}$$

The Keplerian rotation of the planets in our Solar system corresponds to $\alpha = 0$: virtually all mass is located in the center, and $M(R) = M_0$ for all values of R. In Table 12.1 we summarize a few possible rotation curves.

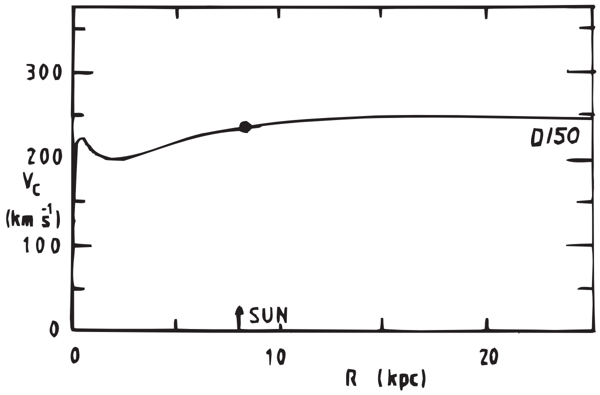

Figure 12.2 The Rotation Curve of the Galaxy.

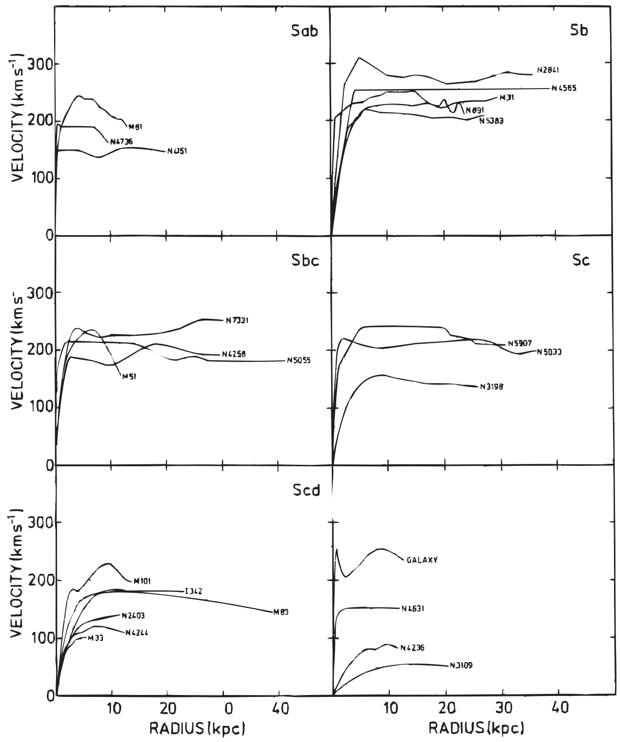

Figure 12.3 Rotation Curves of 25 Galaxies of Different Type, Determined from $\mathrm{H}\,I$ Measurements with the Westerbork Synthesis Radio Telescope. From A. Bosma, 1978 Ph.D. Thesis.

From the empirically determined function $v\,(R)$ one may, in principle, estimate the function $M\,(R)$. As Figures 12.2 and 12.3 show, most galaxies, including our own, have relatively complicated $v\,(R)$ curves. In the very center, $v\,(R)$ usually rises rapidly with R ($\alpha \approx 3$, among others corresponding to $M = M_0 R^3$) suggesting that we are dealing with a homogeneous spherical distribution of stars. With increasing radius R, the rapid rise of $v\,(R)$ flattens off and even may change into a decline ($0 \leq \alpha \leq 1$). This could be interpreted as a drop in mass density just outside the central 'bulge' of the galaxy. A subsequent slow rise ($1 \leq \alpha \leq 2$) suggests that with increasing R, the disk mass density decreases although total mass still increases.

From the Galactocentric distance of the Sun, $R_0 = 28000$ light-years, and the Sun's orbital velocity $v_o\,(R_o) = 220\mathrm{km\ s}^{-1}$ we may estimate that the *Galactic mass* is $M \approx 2 \times 10^{11}\,\mathrm{M}_\odot$. The interstellar gas ($\mathrm{H}\,I$, H_2, He) represent only 5 per cent of that amount. The rotation curves of most other large galaxies are similar. It is very remarkable that these $\mathrm{H}\,I$-derived rotation curves do not turn over to a Keplerian decline at large radii. Even at very large R, sometimes extending several times beyond the optical boundaries, $\alpha \approx 1$.

α	$M(R)$	$v(R)$	Rotation curve	Remarks
0	M_0	cst $R^{-1/2}$	Descending	Keplerian; Solar System
1	$M_0 R$	cst	Flat	
2	$M_0 R^2$	cst $R^{+1/2}$	Ascending	Constant Density
3	$M_0 R^3$	cst R	Ascending	Solid-Body Rotation (v/R is constant)

Table 12.1 Rotation Curve examples.

This leads to the inescapable conclusion that these galaxies have large amounts of matter with much mass, but little or no light. This nonluminous mass cannot be due to stars, nor to gas or dust. The nature of this 'missing mass', or *dark matter* is one of the greatest mysteries of modern astronomy.

12.3 The Galactic Center

Although the very center of the Galaxy suffers from a visual optical depth $\tau_V = 25$, the strong wavelength dependence of interstellar extinction makes the obscuring dust largely or completely transparent at infrared and radio wavelengths. Technical developments in the last few decades have made possible the construction of telescopes and detectors operating at these wavelengths with ever greater resolution and sensitivity. With the most modern telescopes in the world (ESO's VLT in Chile and California's Keck Telescope in Hawaii) individual stars in the very center of the Galaxy can now be seen at infrared wavelengths. Near the center, stellar densities are very high, of the order of 10^5 $M_\odot$ per cubic light-year. These stars experience close encounters so frequently (every million years) that inside a radius $R \leq 650$ lightyears, they behave as particles in a gas with an isothermal velocity distribution. However, within a radius $R \leq 10$ light-years the stars move too rapidly. This is explained by the presence of an additional, very compact mass at the center. Since 1992, astronomers have been tracking the orbits of a few stars very close to the Galactic nucleus. They have established that the Galactic nucleus is a massive *black hole* with $M_0 = 2.6 \times 10^6$ $M_\odot$. From Section 9.7 we find that it should have a of radius of 0.05 Astronomical Units, or only ten times the radius of the Sun! At present, it does not appear to accrete much, and it does not produce much energy. In the distant past, this may have been quite different.

Chapter 13
Galaxies and Clusters

13.1 Galaxy Classification

Based on their apparent morphology, galaxies are classified as elliptical galaxies, spiral galaxies and irregular galaxies. Additional classification criteria are the degree of ellipticity, the presence of a central stellar 'bar' or a central stellar ring, the relative apperance of bulge and disk, or peculiarities in galaxy morphology. There are several classification systems but none of these is entirely satisfactory. One might even doubt whether a 'normal' galaxy exists.

- **Elliptical galaxies (E)** have the form of triaxial ellipsoids. They contain mostly (old) stars and have little or no interstellar gas and dust. All gas originally present appears to have been converted into stars. Some elliptical galaxies still contain some gas and dust, but that appears to have fallen in more recently. The largest galaxies known are invariably (giant) ellipticals, but dwarf ellipticals also exist.
- **Spiral galaxies (S)** have a central bulge of relatively old stars surrounded by a disk with many young stars. The central bulge resembles a (small) elliptical galaxy. Spiral galaxies with dominating central bulges are classified *Sa*. Galaxies of type *Sb* have smaller bulges and in galaxies of type *Sc* the disk is dominant. Sa galaxies are poor in interstellar gas, whereas Sc galaxies are rich in gas and dust. Sa and Sb galaxies usually have two major spiral arms, while the disk of a Sc galaxy is often rather cluttered.
- **Irregular galaxies (Ir)** generally are subdivided into galaxies resembling the Large Magellanic Cloud (size > 10000 lightyears, mass $> 10^8 - 10^9\ M_\odot$) and smaller dwarf irregulars, usually no larger than 5000 lightyears. They contain large amounts of gas.
- **Ellipsoidal galaxies** are primarily known in the Local Group. They are small, contain no gas or dust, show no clear structure and are rather faint. They may represent a pathological class of dwarf galaxies severely distorted by tidal encounters.

13.2 Clusters

There are many more galaxies in the Universe than our own. Inside galaxies, space is almost empty. Stars having characteristic sizes of 10^9 m are separated by distances of the order of 5×10^{16} m, typically 50 million times their own size! In contrast, galaxies themselves occur in groups or clusters. In a cluster, galaxies are separated by only a few dozen to perhaps a hundred times their diameter. The clusters themselves are separated by distances only ten times their own diameter. Thus, the Universe is relatively densely populated with galaxies.

The great majority of galaxies belong to a *group or cluster*. The Milky Way Galaxy is part of a small cluster, which we call the Local Group (diameter about 4 million light-years). The Local Group itself is a satellite of another well-known, very rich cluster in the constellation Virgo (diameter about 15 million light-years. The heart of this Virgo Cluster is at a distance of 45 million light-years. Thousands of clusters are known. Many of these clusters may themselves form even larger entities called *superclusters*. The Local Group could be considered as an outlying member of the Virgo supercluster. Whether or not superclusters in turn form metaclusters is at present speculative. However, it is clear that we live in a very hierarchic Universe.

In clusters, the most massive galaxies are almost always found at the center. Although not all clusters are relaxated systems (yet), mutual interactions, working towards equipartition of energy, generally have brought these massive systems to the cluster center. Almost always, the most massive central galaxies in large cluster are giant ellipticals.

13.3 The Local Group

After the Milky Way Galaxy, the best known member of the Local Group is the Andromeda Galaxy (M 31) which in winter is just visible with the naked eye in the constellation Andromeda. M 31 is a giant spiral galaxy (transitional type Sab), in size and mass comparable to the Milky Way Galaxy (transitional type Sbc). It is seen almost from the side and it is at a distance of nearly 2.5 million lightyears. The Andromeda Galaxy is accompanied by two elliptical dwarf companions, M 32 and NGC 205. It has a large and very bright central bulge and is also surrounded by a large number of globular clusters. The Andromeda Galaxy and the Milky Way are approaching one another with a relative velocity of 120 km/s. In about 5 billion years, they will probably collide.
At about the same distance as the Andromeda Galaxy a much smaller spiral galaxy known as M 33 (dwarf Sc) can be seen in the insignificant constellation Triangulum. It is the third-ranking of the Local Group which contains, in addition, several dozen dwarf galaxies. The nearest and best-known of these are Large Magellanic Cloud and the Small Magellanic Cloud, in the Southern Hemisphere visible to the naked eye. In fact, the Magellanic Clouds are satellites of our own Galaxy. With sizes of only $24000\ \mathrm{ly}$ and $7000\ \mathrm{ly}$ respectively, they are much smaller than our Galaxy with its $55000\ \mathrm{ly}$ diamter. They have masses of only $6 \times 10^9\ M_\odot$ and $1.5 \times 10^9\ M_\odot$, i.e. roughly a per cent of the mass of our Galaxy but are much richer in interstellar gas (10 and 30% respectively by mass). In the Magellanic Clouds, most elements abundances are much lower than in our Galaxy, suggesting much lower past rates of star formation. At distances of about 200000 light-years, the Magellanic Clouds are very close and ideal objects for detailed study. Astronomers were excited when in 1987 a supernova explosion took place in the Large Magellanic Cloud. It was visible to the naked eye for months. The expanding supernova remnant is still being observed reguarly, for instance with the Hubble Space Telescope.

The Magellanic Clouds were captured by the Milky Way several billion years ago. Now, they periodically approach the Galaxy so closely that their orbits are unstable. Eventually, they will crash into the Galaxy. Already, their original presumably regular structure has been seriously affected by tidal forces exerted by the Galaxy. Especially the Small Magellanic Cloud is well on the way of being torn apart, with a long trail of debris following it in space. The eventual impact will have repercussions for the global structure of the Galaxy. However, stellar collisions will be very rare: from the above we see that stars fill only a minute fraction of space, $\left(5 \times 10^7\right)^{-3}$. A similar event is presently taking place. In 1994, a dwarf galaxy was discovered, torn apart and spread out over many degrees on the sky in the constellation of Sagittarius and in the final stage of merging with the Galaxy. Other dwarf galaxies have collided with the Milky Way in the past. Tidal forces have completely demolished the original structure of these dwarfs, and their stars have been absorbed by the Milky Way. However, actual *stellar encounters* are so rare that practically all stars acquired by the Galaxy have managed to retain their original orbital momentum. With satellites such as ESA's Hipparcos, we can measure the motion of very large numbers of stars very accurately. By plotting these motions in velocity space, groups of stars with very similar discordant velocities stand out as relics of past mergers.

In their motions, they have retained their identity notwithstanding the fact that they are by now spread out over the whole Galaxy. Perhaps all the stars in the Galactic halo have their origin in such merger events.

13.4 Merging Galaxies

In a cluster, individual galaxies move in orbits about their common center of mass. Basics physics dictates that in such systems the most massive members gravitate to the center, and that equipartition of energy gives relatively large velocities to dwarf galaxies. The density of galaxies in cluster centers is sufficiently high to make encounters not uncommon. They may pass at such a close range that they greatly distort each other (interacting galaxies), or even collide. Because gravitational potentials are 'soft' targets, colliding galaxies more often than not merge together. Observations with the US-NL-UK infrared satellite IRAS brought to light a class of galaxies with very high infrared luminosities ('ultraluminous galaxies'), often optically inconspicuous because of the great amounts of dust they contain. Almost all ultraluminous galaxies have turned out to represent the last stages of the merger of two massive galaxies – some of them even still have two nuclei.

In a merging or interacting galaxy, gas clouds and even stars may lose angular momentum, causing them to drop to the galaxy center, and creating a concentration of mass at the center. In fact, astronomers now believe that all giant elliptical galaxies have resulted from repeated mergers. Especially when they became gradually more centrally located in the cluster, they must have captured and assimilated many galaxies that passed too close. In the process, they became ever more massive, but also ever more amorphous as their entropy increased. The complex structure of irregular galaxies identifies them as primitive. The least structured galaxies are ellipticals; they must be the most advanced type. Spiral galaxies, less ordered than irregulars, but more ordered than ellipticals, appear to be intermediate. Very deep images obtained with the Hubble Space Telescope show large numbers of galaxies at distances corresponding to the Universe being only 10% of its current age. Almost all of these galaxies are irregular dwarfs or small and very distorted spirals. Almost no elliptical galaxies are seen. It appears that the first galaxies formed in great numbers were irregulars. Their large numbers and the smaller universe they inhabited caused frequent encounters, allowing first large spiral galaxies, and finally giant ellipticals to come into being.

13.5 Active Galaxies and Central Black Holes

Many galaxies exhibit considerable activity in their very center. This may take the form of intense synchrotron radiation from a very compact source in the nucleus, ejection of radio or X-ray jets initially at relativistic speeds, slowing to thousands of kilometers per second. The most extreme examples are so-called *quasars*, extremely compact galaxy nuclei outshining all of the rest of the galaxy. In fact, it appears that almost all galaxies to some degree have an active nucleus, which is probably a supermassive (mass range $10^6 - 10^8$ $\mathrm{M}_\odot$) black hole. These black holes are thought to have resulted from galaxy mergers. The activity is fueled by energy released when the black hole accretes matter, i.e when stars and gas clouds fall in. During this process, half of the potential energy released is radiated away. From Section 8.1 we find that the potential energy of a mass m falling into a black hole of mass M is $E_{grav} = G\,m\,\mathrm{M}/r$, so that the radiated energy is:

$$E = 1/2\;(G\,m\,\mathrm{M}/r_S \;-\; G\,m\,\mathrm{M}/r\,).$$

Because $r \;>>\; r_S$, we may drop the second part of the right-hand side; also, we recall from Section 9.7 that $r_S = 2\;G\;M/c^2$, so that:

$$E = 1/4\;m\,c^2 \tag{13.1}$$

for the duration of the event giving rise to a luminosity:

$$L = 1/4 \ \mathrm{d}m/\mathrm{d}t \, c^2 \tag{13.2}$$

Note that the luminosity has nothing to do with the mass of the central black hole, but is only determined by the *accretion rate* $\mathrm{d}m/\mathrm{d}t$. Thus, quasars probably contain black holes that are rapidly accreting. So-called Seyfert nuclei are also impressive, but much less dominant; they will also contain accreting nuclear black holes, but at substantially reduced rates. Compared to these, the acitivity in the center of our Galaxy is insignificant; the black hole is no longer accreting at any significant rate. Was it more energetic in the past? Has it been dormant for a while only to erupt again? That would truely be a cause for concern. A quasar- or even Seyfert-type outburst could mean the end of life on Earth.

13.6 Missing Mass

The space between the galaxies of a cluster is not empty. With X-ray satellites (NASA-Chandra, ESA-XMM Newton) astronomers observe X-ray emitting gas extending throughout clusters of galaxies. These X-rays are free-free emission by a very hot ($T = 10^5 - 10^6$ K) completely ionized. Because of its high degree of ionization this gas cools only very slowly. With radio telescopes, such as the WSRT, astronomers have also discovered galaxies trailing a radio wake. Apparently, material ejected from these galaxies during their travel through the cluster has been slowed down by intracluster gas and thus marks the orbit of the galaxy over the past tens of millions of years.

By the virial theorem (cf. Section 9.2), we may deduce the total mass of a cluster from the motion of galaxies in the cluster, provided that (a) the cluster is gravitationally *bound*, and that (b) it is *relaxated*. At least for most large clusters this is the case. Very hot intercluster gas has insufficient mass to explain the discrepancy. Analysis invariably leads to the conclusion that clusters are up to ten times more massive than one would estimate from the luminosity of individual galaxies. The problem of missing or dark mass appears here as well, just as it did in the study of galaxy rotation curves.

Chapter 14
The Universe

14.1 Big Bang

The Special Theory of Relativity makes concepts such as distance lose its absolute meaning as it depends on (relative) velocities. As the speed of light c is finite, observing over any distance in space implies observing also over a distance in time. Space and time cannot be separated from each other, and because c is also invariant, we may define invariant distances by: $d^2 = x^2 + y^2 + z^2 = (c\,t)^2$.

Modern cosmology derives from two basic ideas, first formulated in the early 20^{th} century:

(a). *The General Theory of Relativity* starts from the equivalence principle that *gravitational* mass cannot be distinguished from *inertial* mass and considers gravity to be a distortion of space-time.

(b). *The Cosmological Principle* notes that by and large the Universe looks the same in all directions everywhere, i.e. averaged over very large scales, the Universe is homogeneous and isotropic.
Only slightly later, the astronomer Edwin Hubble (after whom the space telescope has been named) discovered that the radiation we receive from all distant galaxies is dopler-shifted to the red. The omnipresent redshift indicates that *all distant galaxies are receding from us* with recession speeds proportional to their distance. A system in which distances uniformly increase with time, is expanding: *we live in an expanding Universe*. Expansion agrees with the Cosmological Principle: any observer, anywhere in the Universe, will see the distant galaxies recede from him. Only on local scales, gravity is strong enough to resist expansion and allows individual galaxies and clusters to remain together. The so-called Hubble constant H_o measures by how much the recession velocity increases per unit distance:

$$v_{rec} = H_o\ d \tag{14.1}$$

This equation may be rewritten to calculate how long ago the expansion started:

$$1/H_o = d\ /v_{rec} \tag{14.2}$$

Thus, the age of the Universe is proportional to $1/H_o$ years. Using powerful telescopes, we may observe galaxies at ever-increasing redshifts, i.e. at ever greater distances and, by implication, ever earlier in the life of the Universe. Because there has not been enough time, light from objects at distances greater than $1/H_o$ light-years has not yet reached us. This distance is therefore called our *Cosmic Horizon*. Because of expansion, the early Universe was much smaller than the present Universe. Since mass and energy are conserved, the early Universe also was hotter and more dense than the present Universe. When it came into being, the volume of space presently within our Cosmic Horizon was infinitesimally small, almost infinitely hot, and almost infinitely dense: *our Universe originated in a 'hot Big Bang'*. This is a bit of a misnomer, as there was no explosion: Space and Time merely appeared simultaneously everywhere. As the Universe, by definition, contains all space and time that we know or can know, it is utterly meaningless to ask what is 'outside', or 'before'. For this very reason, it is also meaningless to ask what 'caused' the Big Bang, if anything did.

14.2 Early Evolution of the Universe

The evolution of the Universe after the Big Bang depended on various properties, such as geometry (closed, flat or open), rate of expansion and content (matter and energy). Its actual age is a function of the present rate of expansion *and* the expansion history (decelerating, constant, accelerating). For instance, if the Universe were to have a critical *mean* density of about 6 H atoms per cubic meter, the expansion would decelerate, without ever becoming zero. Our Universe would be 'flat' and open, i.e. infinite.

From the very beginning the Universe expanded causing its temperature and density to decrease. With decreasing energy-densities, ever-more complex structures could 'freeze out', i.e. were no longer destroyed immediately upon their creation. Quarks, the building blocks of elementary particles, became possible 10^{-8} seconds after the Big Bang. Elementary particles came into being at $t = 10^{-6}$ seconds. It took the Universe, still at temperatures of billions K, about three minutes to fuse protons and neutrons into deuterium ^2eHD and into ^3eHe, and these into helium ^4eHe and some Lithium ^7eLi. This is known as the Big Bang Nucleosynthesis (BBN), and it is worthwhile to recall that even now, hydrogen and helium abundances in the Universe are virtually unchanged from the BBN abundances. It took a few hundred thousand years before atomic nuclei could retain their electrons (i.e. become neutral) and even longer for more fragile molecules to follow.

14.3 Cosmic Background Radiation (CBR)

At the time of the BBN, photons would continually scatter off the free electrons inhabiting the Universe. This multiple scattering caused an ubiquitous and featureless glow because light would be scattered in all directions. Only when protons finally could capture and retain electrons (forming neutral hydrogen) did the space density of free electrons so decrease that the process of multiple photon-scattering abruptly came to an end. From that time on, the universe became transparent, allowing photons to travel almost unhindered through space until the present day. Looking back from our time, we can see through this transparent space until we hit the *Surface of Last Scattering*, beyond which space is filled with a featureless glow through which we can see nothing. Thus, the light of the hot and dense early Universe still reaches us from all directions, and in fact will forever reach us. This *Cosmic Background Radiation*, although originating at high temperatures, has at the present epoch been diluted and cooled to the very low temperature of 2.725 K by the Universal expansion. The CBR, which reaches its peak intensity at a wavelength of about $1\ \mathrm{mm}$, was predicted in 1948 by Gamow, and discovered by accident in 1965 by Penzias and Wilson, who promptly received the 1978 Noble Prize. The CBR has two remarkable properties. First, its spectrum is indistinguishable from *a perfect blackbody Planck spectrum* – perhaps the only such in existence: see Fig. 14.1. Second, the CBR is *extremely smooth*, brightness fluctuations only occur at the 10^{-4} level.

Today, we live in a highly structured and hierarchical universe. Small inhomogeneities or

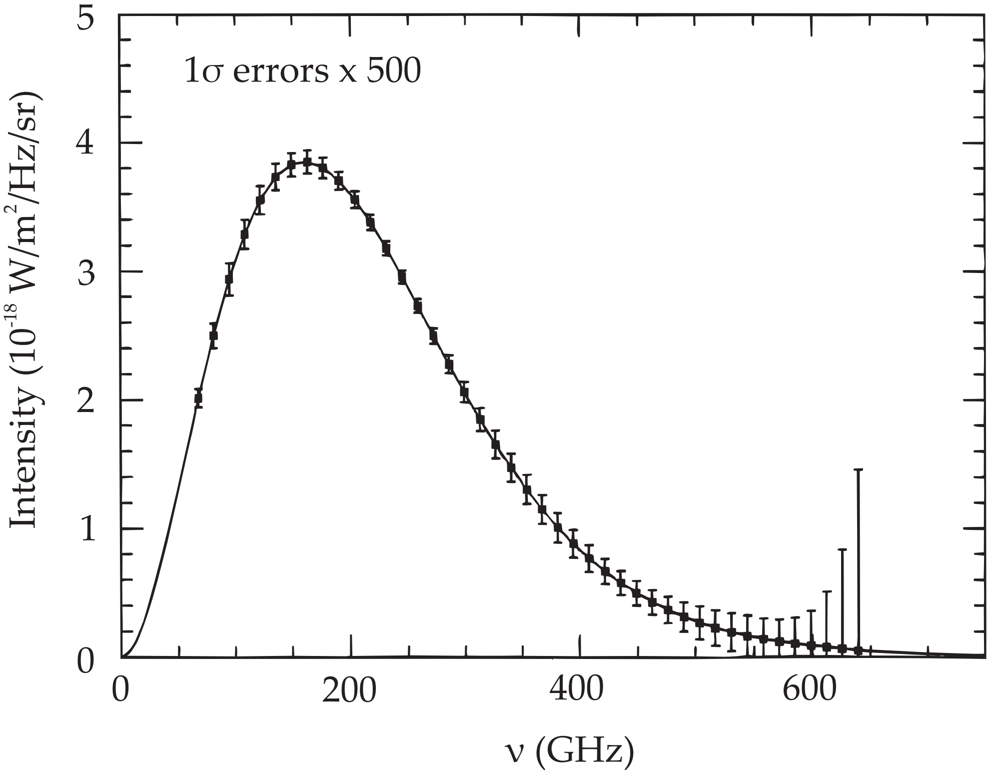

Figure 14.1 Solid line: theoretical Planck curve for T = 2.725 K; vertical bars: measurements of the CBR with uncertainties exaggerated by a factor of 500. Note the perfect fit between experimental data and theoretical model. Courtesy H.J. Habing.

fluctuations must therefore have been present from the very beginning, in time growing to the presently observed Universe (see Table 14.1)

Time Elapsed Since Big Bang	Size of Universe (Present = 1)	Fluctuation Amplitude Local Region of Space
1.2×10^9 yr	0.20	2.0
1.5×10^7 yr	0.01	1.05
5.0×10^5 yr	0.001	1.005

Table 14.1 How fluctuations grew in time to the present Universe.

It appears that in the original very hot, very dense Universe, fluctuations arose by a process called *Cosmic Inflation*. By this process, the Universe experienced one or more periods of extremely rapid expansiuon shortly after the Big Bang.
In order to extract cosmological information from the CBR, its spectrum and its spatial structure must be measured with extremely high accuracy, in particular by eliminating atmospheric effects. This has been done by spacecraft such as COBE (launched 1989) and repeated with ten times higher spatial resolution by WMAP (launched 2001 into the Lagrange point L2 in the Earth-Sun System). The maps produced by WMAP show the minute fluctuations in CBR temperature at the Time of Last Scattering. Statistical analysis of these fluctuations allows determinatoion of various cosmological parameters. For instance, the fact that the fluctuations have a maximum amplitude on scales of one degree means that *the Universe is flat and has a mean density equal to the critical density.*.

The following conclusions have been reached (see for much more explanation and detail: *http://map.gsfc.nasa.gov/index.html*).

- The Universe is infinite and will expand forever.
- The Universe expands at a rate presently given by a Hubble of value $H_o = 71 \pm$

$4\ \mathrm{km\,s^{-1}}$ per megaparsec ($1\ \mathrm{Mpc} = 10^6\ \mathrm{pc} = 3 \times 10^6$ lightyear $= 3.1 \times 10^{22}\ \mathrm{m}$). It does not perceptibly rotate.

- The age of the Universe is $t_o = 13.7 \pm 0.2$ billion years (an unprecedented accuracy of 1%!).
- The Time of Last Scattering was 379 000 years after the Big Bang.
- The first stars started to shine 200 million years after the Big Bang.
- Only 4% *of the Universe's content is luminous matter*; this includes and concludes everything we can see by its radiation. A further 23% *is dark matter*, which can only be observed by its gravity, and of which the nature is largely unknown. However, 73% *of everything is dark energy*, again of unknown nature, and primarily observable by its effect on the expansion of the Universe.
- The Local Group has a velocity of 620 km s^{-1} with respect to the CBR rest frame.

Appendix A
Einstein coefficients

Here we will follow the simple analysis of spectral line formation with which Einstein caused a stir at the beginning of the 20^{th} century.

We consider the simple case of an atomic gas with two energy levels. Level 1 is occupied by n_1 particles (electrons) and has a statistical weight g_1; level 2 corresapondingly has n_2 and g_2. Level 1 is characterized by energy level E, level 2 by energy level $E + h\,\nu_o$. Thermodynamics tells us that at temperature T the equilibrium population of the two levels is given by the Boltzmann partition law:

$$n_1/n_2 = \left(g_1\, e^{-E/k\,T}\right) / \left(g_2\, e^{-[E+h\,\nu_o/k\,T]}\right) \quad = g_1/g_2\, e^{h\,\nu_o/k\,T}$$

A transition from level 1 to level 2 corresponds to absorption of energy, a transition from level 2 to level 1 corresponds to emission. Einstein distinguished three possible processes:

- **Spontaneous emission**: an atom in level 2 spontaneously (no external influence) returns to level 1 emitting the energy difference in the form of a photon at frequency ν_o. Definition: the Einstein A coefficient A_{21} is the probability per second for such a spontaneous transition (unit: s^{-1}).
- **Absorption**: an atom in level 1 encounters and absorbs a photon at frequency ν_o (energy $dE = h\,\nu_o$) and jumps from level 1 to level 2. The probability of this transition depends not only on the properties of the atom, but also on the availability of photons, i.e. the mean intensity I of the incoming radiation at frequency ν_o. Definition: the (first) Einstein B coefficient $B_{12}\,I$ is the transition probability per second for absorption.
- **Stimulated Emission**: there exits a second process proportional to the mean *incoming* radiation intensity I which leads to emission. Definition: the (second) Einstein B coefficient $B_{21}\,I$ is the transition probability per second for stimulated emission. The principle of the maser and the laser (= microwave c.q. light amplification by stimulated emission of radiation) is based on the second Einstein coefficient.

In a gas in thermodynamic equilibrium there are per second as many transitions from level 1 to level 2 as from level 2 to level 1:

$$n_1\, B_{12}\, I = n_2\, A_{21} + n_2\, B_{21}\, I\,,$$

rewrite:

$$I = (A_{21}/B_{21}) \,/\, [(n_1/n_2)\,(B_{12}/B_{21}) - 1]\ .$$

Combine this with the Boltzmann partition law:

$$I = (A_{21}/B_{21}) \,/\, \left[(g_1\, B_{12}/g_2\, B_{21})\ e^{(h\,\nu o/k\,T)} - 1\right]$$

Thermodynamic equilibrium implies thermal equilibrium, so that $I = B$. The above equation must therefore be identical to the Planck function for all temperatures T. One can easily verify that this will yield the Einstein relations:

$$g_1\, B_{12} = g_2\, B_{21}\,,$$

and:

$$A_{21} = 2\,h\,\nu^3/c^2\,B_{21}\,.$$

These equations relate the *atomic* properties A_{21}, B_{21} and B_{12} to each other; they are independent of temperature T. They form, in fact, an extension of Kirchoff's law as they take into account the contribution of non-thermal emission (B_{21}) in a system that is itself in thermal equilibrium (but not necessarily in equilibrium with the incoming radiation field I). It is of great practical importance that one needs to determine no more than just one of the three coefficients to know the other two. Model calculations performed in astrophysics would be impossible without this concept. Another very important asect of the Einstein coefficients is that they do not depend on temperature. This way they may be determined in a laboratory at $T = 300$ K, and subsequently be used for the Sun at temperatures $T = 5800$ K and higher.

Without going into details, we also note there is a relation between the Einstein coefficients (properties of individual atoms) and the source function S (property of the gas, i.e. the ensemble of all atoms):

$$S_\nu = (n_2\,A_{21})\,/\,(n_1\,B_{12} - n_2\,B_{21})$$

Concluding: the Einstein coefficients A and B, which describe the transition probability for transitions between different energy levels of different atoms, can be measured in terrestrial laboratories at arbitrary temperatures. On Earth, one can also determine the statistical weight of the energy levels. Once these are known, one can determine the relative particle density of the level as a function of actual T. In turn, one can also determine the relative intensity of the corresponding spectral lines as a function of T. This is possible for any and every element.

Conversely, one may also determine the temperature of a gas by measuring intensity ratios of different spectral lines of a certain element contained in that gas. Knowing this temperature, one may calculate the particle distribution over the energy levels giving rise to the formation of the observed spectral line. The line intensity is, in principle, proportional to the total number of particles in those energy levels: the more particles, the more intense the line.

Appendix B
Nonthermal Radiation

All radiation with intensities not fully dependent on temperature is called nonthermal, for example stimulated (maser or laser) emission. By some process, an atomic energy level is overpopulated with respect to the population representing thermal equilibrium. The intensity of radiation produced when electrons spontaneously fall back to lower energy levels is proportional to the number of electrons in the excited level, and therefore not dependent on temperature.
In astronomical plasmas (fully ionized gases) a variety of radiative processes may lead to the production of nonthermal radiation. A frequently occurring process produces synchrotron radiation, which was first discovered on Earth in powerful particle accelarators (synchrotrons).

B.1 Synchrotron Radiation

Measurements show that the entire Galaxy is filled with diffuse radio radiation, and that radiative intensities rapidly decrease with increasing frequency:

$$F_\nu \propto \nu^\alpha \, ;$$

a representative value is $\alpha = -0.75$. As thermal free-free transitions produce radio emission with intensities that are almost frequency-independent ($\alpha \approx -0.1$), the observed radiation must be nonthermal in nature. This synchrotron radiation is generated when relativistic (speeds close to c) electrons are injected into a magnetic field. Being electrically charged, they will start to spiral around the field lines and in the process loose energy in the form of photons. The energy of the electrons is very high. For $v \approx c$, the energy of a single electron is:

$$E_\mathrm{e} = mc^2 / \left(1 - v^2/c^2\right)^{1/2} \, ,$$

where $mc^2 = 5 \times 10^5 \text{ eV} = 8 \times 10^{-7} \text{ erg} = 8 \times 10^{-14} \text{ W}$.

The spiralling orbit of the electron has a radius:

$$r_H = E/\,\mathrm{e}\,H = 3.3 \times 10^{-3} E_\mathrm{eV}/\,\mathrm{e}\,H \text{ cm} \, , \qquad \text{(Larmor radius)}$$

where e is the charge of the electron, and H is the magnetic field strength in Gauss. In the Galaxy, one typically finds $H = 5 \times 10^{-6}$ Gauss and $E_\mathrm{eV} = 10^9$ eV, resulting in a characteristic radius $r_H \sim 5 \times 10^9$ m (0.03 AU). A relativistic electron emits photons at a frequency of:

$$\nu_c = (3/4\pi)\,(\mathrm{e}/mc)\,(H \sin\theta)\,\left(E/mc^2\right)^2 \, ,$$

where m is the mass of the electron and θ the angle between the magnetic field and the direction of motion of the electron.

The energy spectrum of all relativistic electrons together (also known as the injection spectrum) is usually written as:

$$N(E) = k\,E^{-\gamma}\,,$$

where $N(E)$ is the number of electrons in the small energy range dE around energy E, and k a constant. By summing the radiation emitted by all electrons at all energy levels, we obtain continuum spectrum:

$$F_\nu = \mathrm{cst}\,(H\sin\theta)^{(\gamma+1)/2}\,\nu^{(-\gamma+1)/2}\,.$$

In the Milky Way, we observed $F_\nu \propto f(H)\,\nu^{-0.75}$, from which we deduce an injection spectrum with $\gamma = 2.5$. As they are constantly losing energy, relativistic electrons have a limited life span. For a relativistic electron with energy E, this is determined by its energy loss per unit time:

$$dE_{\mathrm{GeV}}/dt = 3.8\times 10^{-6}\,(H\sin\theta)^2\,E^2_{\mathrm{GeV}}\ \mathrm{GeV}\,/\,\mathrm{s}\,. \qquad (1\ \mathrm{GeV} = 10^9\ \mathrm{eV})$$

The half-life of (existing) 8 GeV electrons in the Galaxy is approx. 4×10^7 year; that of a 1 GeV electron is eight times longer. Both life spans are, however, very much shorter than the age of the Galaxy (about 10^{10} year).

B.2 Origin of Relativistic Electrons

Thus, in our Milky Way (and in other galaxies) new relativistic electrons must constantly be injected. Such electrons originate in supernova remnants, rapidly expanding gas masses with strong magnetic fields that form the aftermath of a supernova explosion. Supernova remnants sweep up pre-existing interstellar gas, and also produce shock-waves in the interstellar medium. In these shockwaves, non-relativistic electrons can be accelerated almost to the speed of light. Fast-rotating neutron stars with their strong magnetic field are also a source of relativistic electrons.

Young supernova remnants often have a considerably more flat radio spectrum with a *spectral index* $\alpha = -0.3$, implying an injection spectrum with $\gamma = 1.6$. Their radio spectrum is deceptively similar to a thermal spectrum but has nothing to do with it. In supernova remnants, electrons with very high energies are relatively more common than in the Galaxy as a whole. The diffusion process, by which relativistic electrons move into the Galaxy away from their place of birth, takes time. In addition, electrons with high energies lose their energy more quickly than electrons with lower energies. Thus, the synchrotron radiation of the Milky Way is produced by electrons with energy distribution in which the higher energies are poorly represented, resulting in a steep radiation spectrum.

Well-known supernova remnants are the Crab nebula (created by a stellar explosion in 1054 AD), Tycho (1572), Kepler (1604) and Cassiopei A (approx. 1700). The Crab nebula is an optically visible object. Visually, Cas A is unimpressive, quite the opposite of its radio appearance. It is the strongest radio source in the constellation of Cassiopeia (hence its name) and, after the Sun in its active period, the strongest radio source in the sky.

Synchrotron radiation is also released during processes, not fully explained, which occur in the active nuclei of galaxies and in quasars.

Appendix C
Degenerate Matter

Up to very high densities, the ideal gas law $pV = nkT$ remains valid. In such a gas, particles move thermally at greatly differing speeds. The velocity distribution of these particles is described by Maxwell's equation: $f(v) = \mathrm{cst}\, e^{-mv^2/2kT}$. This is true even for stellar interiors where densities $\rho = 10^7$ kg/m^{-3} occur, but no longer at the typical densities $\rho = 10^9$ kg/m^{-3}) found in a white dwarf. These stars are so compact that particles are separated by such small distances that quantum effects come into play.

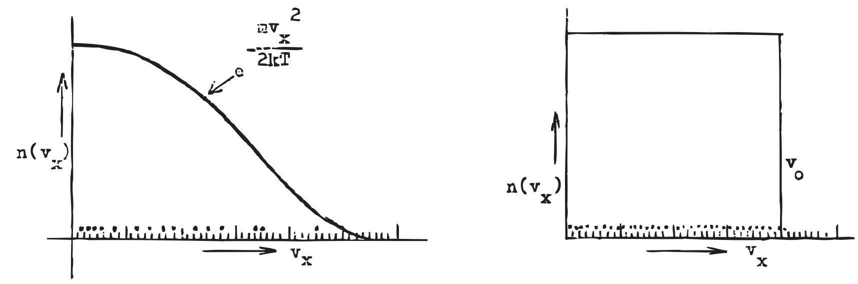

Figure C.1 (left) Maxwellian distribution: the particles are statistically distributed over the quantum states; (right) At very high densities all quantum states are filled: the velocity distribution degenerates. From: J. van der Rijst & C. Zwaan 1975 *Astrofysica*

Two conditions now gain predominance. *Pauli's exclusion principle* states that no two electrons can be in an identical quantum-mechanical state. Heisenberg's uncertainty principle states that position and momentum of a particle cannot be defined with an accuracy better than Planck's constant for the product of the uncertainties: Δx $\times\Delta \mathrm{p}_x > h$. Combination of these two principles leads to the conclusion that even at $T = 0$ K, free electrons should have movement and thereby exert pressure. The nonthermal quantum-mechanical contribution to the total pressure is known as the electron-degeneracy pressure. At very high densities, the distances between electrons Δx are very small, and Pauli's principle causes the electrons to all have momenta differing by at least $h/\Delta \mathrm{p}_x$ from each other. At the same time, Heisenberg's principle causes $\Delta \mathrm{p}_x$ for small Δx. Thus, in very high-density environments, free electrons must have high speeds $v = p/m_e$, causing a sufficiently high internal pressure to resist collapse even against the extreme self-gravity of a white dwarf. When a white dwarf cools, it does not contract.

If the stellar mass exceeds that appropriate for a white dwarf, increasing compression forces the degenerate electrons onto the atomic nuclei, where they fuse with the protons into neutrons, thus forming a neutron star with characteristic densities of the order of $10^{18}\,\mathrm{kg\ m}^{-3}$. Because the neutrons are almost touching, nuclear forces (i.e. the strong interaction) becomes important. Theories of the internal structure of neutron stars require full application of the nuclear forces as well as the non-Newtonian gravitational forces as described e.g. in Einstein's General Theory. The Special as well as the General Theory of Relativity poses limits to the ability of neutron matter to resist compression. Increasing pressure (i.e. stellar mass) beyond these limits leaves the neutrons no other choice than being crushed. Stellar surface gravities imply escape speeds beyond the speed of light: a black hole is born.

Appendix D
Nuclear Fusion

(Note: following Van der Rijst & Zwaan (1975), this appendix is based on a lecture course at Utrecht University by E.P.J. van den Heuvel)
In the centre of the Sun, the temperature is about 1.8×10^7 K. At this temperature, radiation is mostly (Wien's displacement law) emitted at a wavelength of 0.2 nm. The mean free path of a photon in the Sun's interior is about 20 cm. The energy must therefore work its way out of the Sun in a long series of absorption and re-emission processes. As a consequence, energy released in the Solar centre takes approximately 10^6 years to reach the edge of the Sun.

In the Sun, the cycles shown in Tables D.1 and D.2 all contribute to the production of energy: PP I about 56%, PP II about 41%, PP III about 0.01% and the CNO-cycle 3%. The neutrinos of the PP III cycle have the highest energies and the greatest chance of being detected because the reaction probability with Cl is proportional to E^2,

D.1 The Solar Neutrino Problem

Helium fusion cycles occurring at temperatures $T < 1.5 \times 10^7$ K are given in Table D.1. At temperatures $T > 1.5 \times 10^7$ K, helium fusion does not occur directly from protons, but requires intermediate elements as catalyst. The relevant cycles are given in Table D.2. Part of the energy released by nuclear fusion is taken away by neutrinos ν; for the P III cycle this is 29%. The neutrinos (having no charge and negligible rest mass, but possessing energy and momentum) move right through the Sun (as well as the Earth), as opposed to the photons of thermal radiation).

The Sun converts per second:

$$\frac{4 \times 10^{26}\ \mathrm{J}}{6.5 \times 10^{14}\ \mathrm{J}\,/\,\mathrm{kg}} = 6 \times 10^{11}\ \mathrm{kg}\ ,$$

of its mass, corresponding to 4×10^{37} H nuclei. This means that 2×10^{37} neutrinos are released each second. At the distance $R =$ of the Earth to the Sub, these neutrinos pass through a sphere of surface area $4\pi R^2 = 1.7 \times 10^{23}\ \mathrm{m}^2$ (n). Thus, the number of neutrinos emitted by the Sun passing each second through each m^2 on Earth is:

$$\frac{2 \times 10^{37}}{1.7 \times 10^{23}} \approx 1.2 \times 10^{14}\ .$$

They take away approximately 5% ($2 \times 10^{25}\ \mathrm{J}\,/\,\mathrm{s}$) of the total amount of energy produced.

For many years, attempts have been made to detect these neutrinos using reactions based on the weak interaction forces, such as: $^{37}\mathrm{Cl} + \nu \rightarrow {}^{37}\mathrm{A} + \mathrm{e}^-$; the resulting radioactive Argon can be detected. Detection rates are very low, currently about 1 neutrino per hour!

The Proton-Proton Cycle		Mean Lifetime	Energy Produced (MeV)
PP I	$2\,^1\mathrm{H} + 2\,^1\mathrm{H} \rightarrow 2\,^2\mathrm{D} + 2\beta^+ + 2\nu$	10^{10} yr	2.38
	$2\,^2\mathrm{D} + 2\,^1\mathrm{H} \rightarrow 2\,^3\mathrm{He} + 2\gamma$	6 sec	10.98
	$^3\mathrm{He} + {}^3\mathrm{He} \rightarrow {}^4\mathrm{He} + 2\,^1\mathrm{H}$	106 yr	12.85
net:	$4\,^1\mathrm{H} \rightarrow {}^4\mathrm{He} + 2\beta^+ + 2\nu + 2\gamma$		26.21 (2% ν-losses)
PP II	$^1\mathrm{H} + {}^1\mathrm{H} \rightarrow {}^2\mathrm{D} + \beta^+ + \nu$		1.19
	$^2\mathrm{D} + {}^1\mathrm{H} \rightarrow {}^3\mathrm{He} + \gamma$		5.49
	$^3\mathrm{He} + {}^4\mathrm{He} \rightarrow {}^7\mathrm{Be} + \gamma$		1.58
	$^7\mathrm{Be} + \mathrm{e}^- \rightarrow {}^7\mathrm{Li} + \gamma + \nu$		0.05
	$^7\mathrm{Li} + {}^1\mathrm{H} \rightarrow 2\,^4\mathrm{He}$		17.34
net:	$4\,^1\mathrm{H} \rightarrow {}^4\mathrm{He} + 2\beta^+ + 2\nu + 2\gamma$		25.65 (4% ν-losses)
PP III	$^1\mathrm{H} + {}^1\mathrm{H} \rightarrow {}^2\mathrm{D} + \beta^+ + \nu$		1.19
	$^2\mathrm{D} + {}^1\mathrm{H} \rightarrow {}^3\mathrm{He} + \gamma$		5.49
	$^3\mathrm{He} + {}^4\mathrm{He} \rightarrow {}^7\mathrm{Be} + \gamma$		1.58
	$^7\mathrm{Be} + {}^1\mathrm{H} \rightarrow {}^8\mathrm{B} + \gamma$		0.14
	$^8\mathrm{B} \rightarrow {}^{8x}\mathrm{B} + \beta^+ + \nu$		7.70
	$^{8x}\mathrm{B} \rightarrow {}^4\mathrm{He}$		3.00
net:	$4\,^1\mathrm{H} \rightarrow {}^4\mathrm{He} + 2\beta^+ + 2\nu + 2\gamma$		19.10 (29% ν-losses)

Table D.1 Cycles at Temperatures $T < 1.5 \times 10^7$ K.

The CNO-cycle		mean life	energy production (MeV)
	$^{12}\mathrm{C} + {}^1\mathrm{H} \rightarrow {}^2\mathrm{N} + \gamma$	10^7 years	1.95
	$^{13}\mathrm{N} \rightarrow {}^{13}\mathrm{C} + \beta^+ + \nu$	7 minutes	1.50
	$^{13}\mathrm{C} + {}^1\mathrm{H} \rightarrow {}^{14}\mathrm{N} + \gamma$	3×106 years	7.54
	$^{14}\mathrm{N} + {}^1\mathrm{H} \rightarrow {}^{15}\mathrm{O} + \gamma$	3×106 years	7.35
	$^{15}\mathrm{O} \rightarrow {}^{15}\mathrm{N} + \beta^+ + \nu$	80 seconds	1.73
	$^{15}\mathrm{N} + {}^1\mathrm{H} \rightarrow {}^{12}\mathrm{C} + {}^4\mathrm{He}$	10^5 years	4.96
net:			25.03 (6% ν-losses)
or (one time in a 1000):			
	$^{15}\mathrm{N} + {}^1\mathrm{H} \rightarrow {}^{16}\mathrm{O} + \gamma$		12.11
	$^{16}\mathrm{O} + {}^1\mathrm{H} \rightarrow {}^{17}\mathrm{F} + \gamma$		0.59
	$^{17}\mathrm{F} \rightarrow {}^{17}\mathrm{O} + \beta^+ + \nu$		1.76
	$^{17}\mathrm{O} + {}^1\mathrm{H} \rightarrow {}^{14}\mathrm{N} + {}^4\mathrm{He}$		1.20
net:			15.66
net:	$4\,^1\mathrm{H} \rightarrow {}^4\mathrm{He} + 2\beta^+ + 2\nu + 2\gamma + 26.72$ MeV (including ν-energy)		

Table D.2 Cycles at Temperatures $T > 1.5 \times 10^7$ K.

Surprisingly, the measured rate of neutrinos from the Sun was only one third of the expected flux (the Solar Neutrino Problem). This deficiency of neutrinos has long been a major problem in Solar astrophysics.
Each of the massive leptons (electron, muon and tau) has its associated neutrino. Neutrino's produced by the proton-proton cycles in the Solar interior are all electron neutrino's. The current generation of neutrino detectors can, in principle, distinguish between electron and muon neutrinos. A muon neutrino interacting with a nucleus produces an energetic muon which travels only a short distance, emitting a *sharply outlined cone* of Cerenkov radiation. An electron neutrino interaction produces an energetic electron, which in turn generates a shower of electrons and positrons, each with its own Cerenkov cone. The overlapping Cerenkov cones now produce a *diffuse circle of light*. Tau neutrinos are not detected by these detectors.
Results from the Sudbury Neutrino Observatory, specifically designed to study the Solar Neutrino problem, imply that a fraction of the electron neutrinos produced by the Sun are transformed into muon neutrinos on the way to the Earth ('neutrino oscillations'). The observations at Sudbury are then consistent with the Solar models of neutrino flux. This is still a very hot topic; much of the above has been taken from world-wide-web sites such as:
http://www.maths.qmw.ac.uk/ lms/research/neutrino.html
and other responding to a serach on *solar neutrino problem*.

D.2 Fusion Reaction Rates

The rate at which energy is generated by the fusion reactions discussed is a function of temperature T, density ρ and the weight percentages of the reacting elements, X_A and X_B. The corresponding molecular weights of the elements are μ_A and μ_B, so that the number of particles per unit volume of each of these elements is:

$$N_A = \frac{X_A \rho}{\mu_A} ,$$
$$N_B = \frac{X_B \rho}{\mu_B} ,$$

Reaction rates will be proportional to:

1. the collision probability of the two types of elements;
2. the energy of the colliding particles.

Obviously, the second factor will depend on the temperature of the gas, expressed as a function $f(T)$. The first factor, the collision probability, is proportional to:

$$N_A N_B = \frac{X_A X_B \rho}{\mu_A \mu_B} .$$

Thus, the amount of energy generated per second per unit volume will be:

$$\epsilon_{AB} = \frac{X_A X_B \rho^2}{\mu_A \mu_B} f(T) .$$

N.B.: here and in the following, energy-generation rates (energy yields) are in unit volume. If instead values per unit mass are required, one should muliply by the volume of a mass unit of mass, i.e. the density ρ.
It appears that $f(T)$ can be approximated numerically quite well by:

$$f(T) = kT^{\nu} ,$$

where the PP-cycles have $\nu \approx 4$ and the CNO-cycle has $\nu = 12$. The constant k depends on the reaction considered.

The triple-alpha process	energy production (MeV)
${}^4\mathrm{He} + {}^4\mathrm{He} \to {}^2\mathrm{Be} + \gamma$	-0.095
${}^8\mathrm{Be} + {}^4\mathrm{He} \to {}^{12}\mathrm{C} + \gamma$	7.400
net: $3\,{}^4\mathrm{He} \to {}^{12}\mathrm{C} + 2\gamma$	7.305

Table D.3 The triple-alpha process.

We thus expect for the PP-cycles:

$$\epsilon_{PP} = X_{\mathrm{H}}^2 \rho^2 T^4 \,,$$

and for the CNO-cycle:

$$\epsilon_{CNO} = X_{\mathrm{H}} X_{\mathrm{C}} \rho^2 T^{12} \,.$$

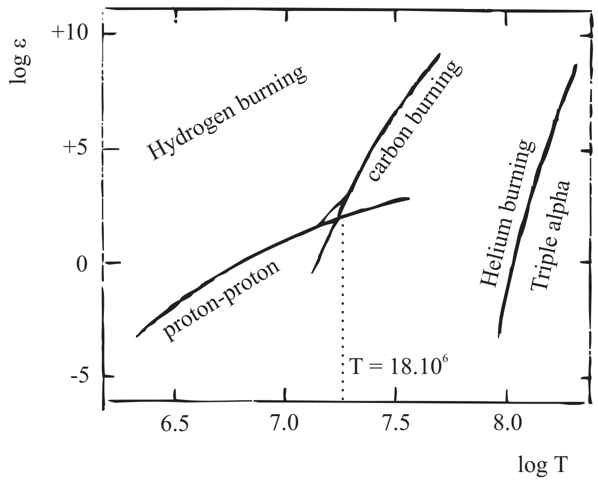

Figure D.1 The energy yield ϵ versus temperature T for $\rho X_{\mathrm{H}}^2 = 100$ and $X_{\mathrm{C}} = 0.005 X_{\mathrm{H}}$. Below $T = 18 \times 10^6$ K the proton-proton cycle dominates, above $T = 18 \times 10^6$ K the CNO-cycle.

D.3 The Triple-Alpha Process

In addition to the processes discussed in the previous sections Figure D.1 shows a third process for generating energy. This process only comes into play at temperatures $T > 100 \times 10^6$ K and is called the *triple-alpha process* in which three helium nuclei fuse into one carbon nucleus (see Table D.3).

The ${}^8\mathrm{Be}$ nucleus is unstable and decays into two helium nuclei releasing 95 keV. In a continuous process of ${}^8\mathrm{Be}$ production and decay, an equilibrium concentration of ${}^8\mathrm{Be}$ is always present. By capturing alpha (${}^4\mathrm{He}$) particles they further fuse to ${}^{12}\mathrm{C}$ and release additional energy.

As before, the energy yield $\epsilon_{3\alpha}$ is proportional to $\rho^3 Y^3$ (Y = helium weight fraction) and a power of temperature, T^ν:

$$\epsilon_{3\alpha} = \mathrm{cst}\, \rho^3 Y^3 T^{30} \,.$$