The Interstellar Medium

Prof. dr. Alex DE KOTER

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

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

Version 1.46.00

Physical Constants

Name	Symbol	CGS Value	CGS units
Speed of light in a vacuum	c	2.998×10^{10}	$\mathrm{cm\,s^{-1}}$
Planck constant	h	6.626×10^{-27}	erg s
	$\hbar$	1.055×10^{-27}	erg s
Gravitational constant	G	6.673×10^{-8}	$\mathrm{cm^3\,g^{-1}\,s^{-2}}$
Atomic mass unit	$\mathrm{m_{amu}}$	1.661×10^{-24}	g
Mass of hydrogen	$\mathrm{m_H}$	1.673×10^{-24}	g
Mass of proton	$\mathrm{m_p}$	1.673×10^{-24}	g
Mass of neutron	$\mathrm{m_n}$	1.675×10^{-24}	g
Mass of electron	$\mathrm{m_e}$	9.109×10^{-28}	g
Avagadro's number	N_A	6.022×10^{23}	
Boltzmann constant	k	1.381×10^{-16}	$\mathrm{erg\,K^{-1}}$
Radiation density constant	a	7.565×10^{-15}	$\mathrm{erg\,cm^{-3}\,K^{-4}}$
Stefan-Boltzmann constant	σ	5.671×10^{-5}	$\mathrm{erg\,cm^{-2}\,K^{-4}}$
Fine structure constant	α	7.297×10^{-3}	
Bohr magneton	μ_B	9.274×10^{-21}	$\mathrm{erg\,G^{-1}}$
Rydberg constant	R_∞	2.180×10^{11}	erg

Astronomical Constants

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

Conversion Factors

Name	Symbol	Value	CGS units
Year	yr	3.156×10^7	s
Arcsec	”	4.848×10^{-6}	radians
Electron volt	eV	1.602×10^{-12}	erg
Solar mass per year	$M_\odot \mathrm{yr}^{-1}$	6.303×10^{25}	$\mathrm{g\,s^{-1}}$

Contents

1

Introduction to interstellar matter

Though the mass of most galaxies is primarily in the form of dark matter particles, it is the baryons – accounting for perhaps $\sim$10 percent of the total mass – that determine the visible appearance of galaxies. At early times, the baryonic mass in galaxies was primarily in the gas in the interstellar medium (ISM). As galaxies evolve, the interstellar medium is gradually converted to stars and stellar remnants, and some part of the ISM may be ejected from the galaxy in the form of galactic winds. Infalling gas from the intergalactic medium may add to the mass of the ISM. The mass flow of the baryons in our galaxy is schematically shown in Fig. 1.1.

Donald Osterbrock states that 'the interstellar medium is anything not in stars'. If that is so, what then are the constituents of the ISM. We may distinguish (see Draine 2011):

- Interstellar gas: Ions, atoms, and molecules in the gas phase, with velocity distributions that are very nearly thermal.
- Interstellar dust: Small solid particles, mainly less than $\sim$1 μm in size, mixed with the interstellar gas.
- Cosmic rays: Ions and electrons with kinetic energies far greater than thermal, often extremely relativistic.
- Electromagnetic radiation: Photons from many sources, including the cosmic microwave background (CMB); stellar photons; radiation emitted by interstellar atoms, ions, and molecules; thermal emission from interstellar grains that have been heated by starlight; free-free emission from interstellar plasma; synchrotron radiation from relativistic electrons; and gamma-rays emitted in nuclear transitions and π^0 decays.
- Interstellar magnetic field: The magnetic field resulting from electric currents in the interstellar medium; it guides the cosmic rays, and in some parts of the ISM, the magnetic field is strong enough to be dynamically important.
- The gravitational field: This is due to all of the matter in the galaxy – ISM, stars, stellar

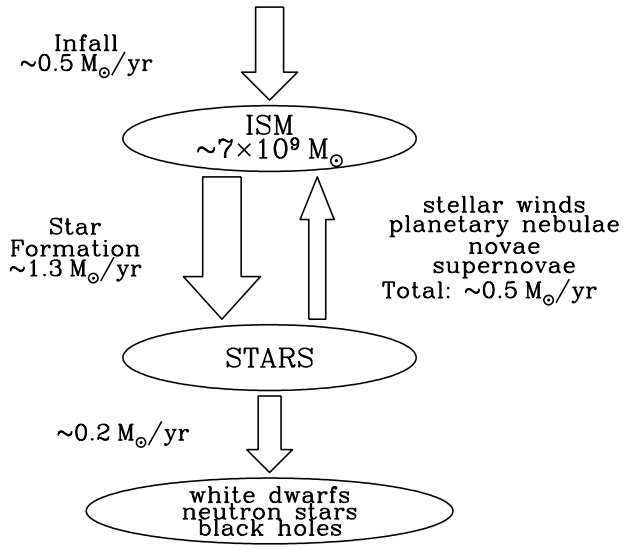


Figure 1.1: *The flow of baryons in the Milky Way. From: Bruce Draine*, Physics of the Interstellar and Intergalactic medium.

remnants, and dark matter – but in some regions, the contribution of the ISM to the gravitational potential leads to self-gravitating and even collapsing clouds.

- Dark matter: Though gravitationally dominant, it is currently unknown whether, and if so to what extend, dark matter particles interact non-gravitationally with baryons, electrons, or magnetic fields. Nor is it known if they decay or annihilate into particles that interact with baryons, electrons, or magnetic fields.

All these components of the ISM interact and influence each other. In addition, the ISM interacts with stars. Stars form from dense interstellar clouds, and once formed the radiation of stars affects the physical and chemical conditions of the ISM. Stellar winds and supernovae return much of the stellar material back into the ISM, but these stellar ejecta are enriched in metals due to nucleosynthesis. Furthermore, stellar ejecta carry energy and momentum into the ISM. The enriched material mixes in the ISM and eventually will be incorporated in new stars (and planets). This cycle of matter is an important aspect of the evolution of galaxies, and motivates the study of the interstellar medium.

In the Milky Way galaxy today, perhaps 10 percent of the baryons are to be found in the ISM. Interstellar matter is most often found in spiral galaxies such as our own, and (much) less in elliptical galaxies. The gas and dust in the ISM of spiral galaxies is concentrated in clouds along the spiral arms, and from these dense clouds new generations of stars are born. So spiral arms stand out in images because they contain the young, luminous and blue stars (see Fig. 1.2). The hot stars ionize the gas in their surroundings, causing easily visible H II regions.

Elliptical galaxies show much less interstellar gas. Probably most of the gas in these systems has been converted into stars in the past, consequently the star formation rate in these galaxies is much lower than in spiral galaxies.

1.1 Interstellar gas

Interstellar *gas* can have a wide range of temperatures and densities. Because the ISM is dynamic, all densities and temperatures within these ranges can be found somewhere in the Milky Way. However, it is observed that most of the gas particles have temperatures falling close to various characteristic states or thermal phases (see also Table 1.1). In order from coolest to hottest and explicitly referencing the chemical form of the dominant element, hydrogen, we distinguish:

1. Very cold and dense Molecular Clouds (MCs): H_2
   Very cold ($T \sim 10\,\mathrm{K}$) and dense (particle number densities $n > 300\,\mathrm{cm}^{-3}$) molecular gas, distributed in *molecular clouds*. In our galaxy Molecular Clouds comprise ∼20 percent of the baryonic mass of the ISM, but occupy only ∼0.01 percent of its volume. It is on account of the (relatively speaking) high densities in these clouds that molecules can actually form, hence the name. More specifically, it is firstly because the processes forming molecules in interstellar space – primarily two-body gas-phase reactions and catalysis on the surface of dust grains – proceed faster at higher density, and secondly because dust effectively shields the interior of a dense cloud from the ultraviolet radiation which destroys molecules. These clouds are often 'dark' – with visual extinction $A_V \gtrsim 3\,\mathrm{mag}$ through their central regions. The main tracers of cloud structure, dynamics, conditions and composition are millimetre-wavelength molecular emission lines (for instance CO). Dust grains in dark clouds are often coated with mantles of H_2O and other molecular ices. Most molecular clouds are gravitationally bound. The densest regions are likely unstable and sites of new star formation. Note that the gas pressures in the densest regions would qualify as ultra-high vacuum in a terrestrial laboratory.

2. Cold Neutral Medium (CNM): H I, traced through spectral line absorption
   Cold ($T \sim 100\,\mathrm{K}$) gas is distributed in clouds, sheets and filaments of densities $n \sim 30\,\mathrm{cm}^{-3}$ occupying ∼1–4 percent of the ISM. The material is mainly neutral and atomic (though species with low ionization potentials may be ionized). The densities are too low for molecules to form abundantly. These clouds are typically detected through the absorption of background stellar light in optical (e.g. Ca II, Na I and K I) and near ultraviolet (e.g. Mg II) ground-level resonance lines.

   When densities and column densities are sufficiently large, H_2 may form and survive due to H_2 self-shielding. We refer to such clouds as diffuse molecular clouds. They are similar to the CNM, though temperatures may be somewhat lower ($\sim 50\,\mathrm{K}$) and densities somewhat higher ($\sim 100\,\mathrm{cm}^{-3}$). We do not rank them as a separate phase (but

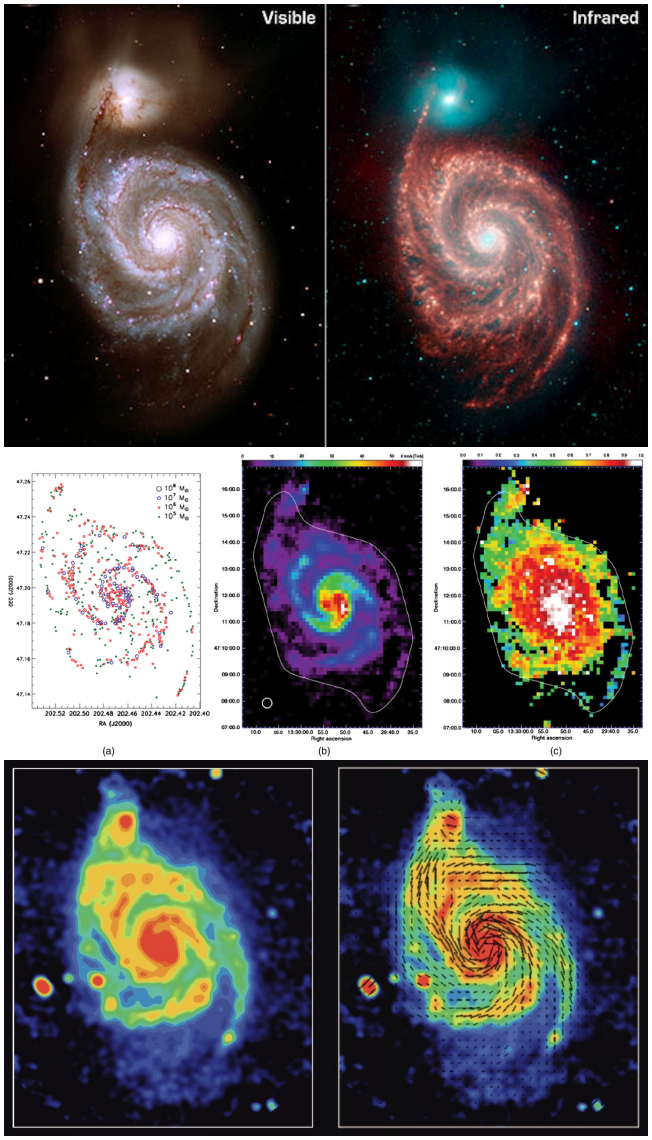


Figure 1.2: *M51a (NGC 5194, the Whirlpool galaxy) and its companion M51b (NGC 5195). Top panels: Visible-light image, with conspicuous dark dust lanes located along the spiral arms, and bright blue star forming regions. Near-infrared image in false blue (3.6 μm) and false red (8 μm) colors. The dust lanes, dark in the left panel, are glowing in 8 μm PAH emission. Note the numerous 'holes', where PAH emission is weak. (From: NASA/JPL-Caltech / R. Kennicutt). Middle panels, from left to right: Location of giant molecular clouds (GMCs) with masses $> 4 \times 10^5\ M_\odot$ (detected in CO 1–0). CO 1–0 line intensity image, smoothed over a 22" beam (circle in lower left corner); violet is 5 $\mathrm{K\,km\,s^{-1}}$, red = 50 $\mathrm{K\,km\,s^{-1}}$. H_2 fraction $2N(H_2)$/ $[2N(H_2) + N(\mathrm{H\,I})]$: green = 0.5, white = 1 (from Koda et al. 2009). Bottom panels, from left to right: 5 GHz emission, mainly synchrotron radiation from relativistic electrons. Magnetic field polarization of 5 GHz emission, revealing that the magnetic field is aligned with the spiral arms. (From NRAO/AUI/NSF / R. Beck and C. Horellou).*

some do). They can be observed with the same diagnostics as the CNM and in addition by CO 2.6-mm emission.

3. Warm Neutral Medium (WNM): H I, traced through spectral line emission
   Warm neutral atomic hydrogen occupies ∼20–40 percent of the volume of the ISM and is located mainly in the boundary layers between molecular clouds and H II regions. It has characteristic temperatures of 5000–8000 K and densities of $\sim 0.5\,\mathrm{cm}^{-3}$, and is traced by the H I emission line at 21 cm. These boundary layers are characterized by the photo-dissociation of the H_2 molecule and are therefore termed photodissociation regions. Some older literature refers to the Warm Neutral Medium as 'warm inter-cloud medium'.

4. Warm Ionized Medium: H II, traced through spectral line emission
   Extended and diffuse gas with temperatures of 6000-10000 K, and densities $\sim 0.3\,\mathrm{cm}^{-3}$ occupying about 15 percent of the volume of the ISM. While primarily photo-ionized (to do so requires about $1/6^{\mathrm{th}}$ of all the ionizing photons emitted by the Galaxy's O and B stars), there is some evidence of shock or collisional ionization high above the plane of the Galaxy. The diffuse H II is traced by low-surface brightness $\mathrm{H}\alpha\ \lambda 6563$ Å emission.

   Nearly 90 percent of the H II in the Galaxy resides in the Warm Ionized Medium, with the remaining 10% in the bright high-density (compact, or at least localized) H II regions that occupy a tiny fraction of the ISM and that arise from the ultraviolet radiation of nearby, hot massive stars. In compact H II regions densities may increase to $\sim 10^4\,\mathrm{cm}^{-3}$. The Orion Nebula, measuring a few pc across, is an example of a bright H II region. Its lifetime is essentially that of the ionizing stars, so 3–10 Myr.

   In addition to the H II regions, photo-ionized gas is also found in *planetary nebulae* – these are created when rapid mass loss during the late stages of evolution of stars with initial mass $0.8M_\odot < M < 8M_\odot$ exposes the hot stellar core, creating a luminous planetary nebula. Individual planetary nebulae fade away on a $\sim 10^4$ yr time scale.

5. Coronal or Hot Ionized Medium (HIM): X-ray and O IV-VI absorption and emission
   Hot, low-density gas heated by shock waves from supernovae, with temperatures exceeding $> 3 \times 10^6$ K and very low densities of $< 0.004\,\mathrm{cm}^{-3}$, occupies ∼50 percent of the ISM. The vertical scale height of this gas is ∼3 kpc, so it is sometimes referred to in the literature as the hot 'corona' of the galaxy. This hot gas is often buoyant and appears as bubbles, with characteristic dimensions of ∼20 pc, and fountains high above and below the disk. The coronal gas volumes may be connected to other coronal gas volumes. The gas is collisionally ionized and cools on ∼Myr timescales. Its primary tracers are absorption lines seen towards hot stars in the far-ultraviolet (e.g. O IV, N V, and C IV) in gas with $T \sim 10^5$ K, and diffuse soft X-ray emission from gas hotter than 10^6 K. In the hottest coronal gas O VI is present.

To this list of phases of interstellar gas, some add a 6th phase, namely gas in stellar outflows. Evolved cool stars lose gas at rates up to a few times $10^{-4}\ M_\odot \mathrm{yr}^{-1}$. This gas streams into

Phase	T (K)	$n_{\rm H}$ (cm s^{-3})	Comments
Molecular clouds H_2 $f_{\rm V} \approx 10^{-4}$ $< n_{\rm H} > f_{\rm V} \approx 0.2$ cm^{-3}	10–50	10^3 – 10^6	Heating by photons from dust Ionization and heating by cosmic rays Often self-gravitating: $p > p$(ambient ISM) Cooling by: – CO line emission – C I fine structure line emission Observed by: – CO 2.6-mm emission – dust far-IR emission
Cool H I (CNM) $f_{\rm V} \approx 0.01$ $< n_{\rm H} > f_{\rm V} \approx 0.3$ cm^{-3}	$\sim$100	30	Heating by photons from dust Ionization by starlight, cosmic rays Pressure equilibrium Cooling by: – Fine structure line emission Observed by: – H I 21-cm emission, absorption – Optical, UV absorption lines
Warm H I (WNM) $f_{\rm V} \approx 0.2 - 0.4$ $< n_{\rm H} > f_{\rm V} \approx 0.2$ cm^{-3}	$\sim$5000–8000	0.5	Heating by photons from dust Ionization by starlight, cosmic rays Pressure equilibrium Cooling by: – Optical line emission – Fine structure line emission Observed by: – H I 21 cm emission, absorption – Optical, UV absorption lines
(Diffuse) H II gas $f_{\rm V} \approx 0.15$ $< n_{\rm H} > f_{\rm V} \approx 0.002$ cm^{-3}	$\sim 10^4$	$0.3 - 10^4$	Heating by photons from O stars Photo-ionized by O stars Either expanding or in pressure equilibrium Cooling by: – Optical line emission – Free-free emission – Fine structure line emission Observed by: – Optical line emission – Thermal radio continuum
Coronal gas (HIM) $f_{\rm V} \approx 0.5$? $< n_{\rm H} > f_{\rm V} \approx 0.002$ cm^{-3}	$\gtrsim 3 \times 10^6$	$\lesssim 0.004$	Shock-heated Collisionally ionized Either expanding or in pressure equilibrium Cooling by: – Adiabatic expansion – X-ray emission Observed by: – UV and X-ray emission – Radio synchrotron emission

Table 1.1: *Phases of the interstellar medium, ordered from coolest to hottest.* $f_{\rm V}$ *is the volume filling factor. Following: Draine,* Physics of the Interstellar and Intergalactic Medium.

Phase	T (K)	$n_{\rm H}$ (cm s^{-3})	Comments
Cool stellar outflows H_2	50–10^3	1 – 10^6	Observed by – CO, OH, and other molecules – dust IR emission
Hot stellar outflows H II	8000 – 30000	0.1 – 10^5	Observed by – UV absorption lines – Hα, He II 4686 emission – Thermal radio continuum

Table 1.2: *Gas in stellar outflows, ordered from coolest to hottest.*

the ISM at velocities up to 30 $\mathrm{km\,s^{-1}}$, leading to relatively high density outflows. Hot stars can have mass-loss rates as high as a few times 10^{-5} $M_\odot \mathrm{yr}^{-1}$ and outflow velocities of several 1000s of $\mathrm{km\,s^{-1}}$. Compared to cool stars, the winds of hot stars tend to be somewhat less dense. The outflowing gas is often referred to as *circumstellar matter* and at some distance from the star will interact with interstellar matter. For completeness, let us complement Table 1.1 with a table for gas in stellar outflows (see Table 1.2).

The co-existence of the different phases of ISM gas

The co-existence of these different phases is in many cases due to the presence of stars, e.g. H II regions surrounding hot stars that are embedded in larger H I regions. The fact that H I and H II regions can exist side by side and that warm and cold regions co-exist needs an explanation. For a typical H I cloud with number density $n = 50\,\mathrm{cm^{-3}}$ and temperature $T = 80\,\mathrm{K}$, the product nT is $2 \times 10^3\,\mathrm{cm^{-3}\,K}$. This value is matched, within a factor of two, by the corresponding product for the warm neutral medium ($n = 0.5\,\mathrm{cm^{-3}}$; $T = 8\,000\,\mathrm{K}$). The numbers suggest that H I clouds and the warm neutral medium co-exist in pressure equilibrium. Indeed, an analysis of the processes that heat and cool the gas point to the cool and warm neutral medium as representing thermally stable states. In a thermally stable state the internal pressure increases when the gas is compressed.

H II regions that have a higher pressure than the surrounding medium, therefore expand, and very dense cores of molecular clouds that are self-gravitating may not be in pressure equilibrium, and collapse.

Large interstellar molecules and Diffuse Interstellar Bands

The interstellar medium contains a population of large molecules, that reveal themselves by conspicuous emission features at 3.3, 6.2, 7.7, 8.6, 11.3, and 12.7 μm. These emission features are characteristic for polycyclic aromatic hydrocarbon (PAH) molecules. PAH molecules are planar structures consisting of carbon atoms organized into hexagonal rings, so-called aromatic rings, with hydrogen attached at the boundary.

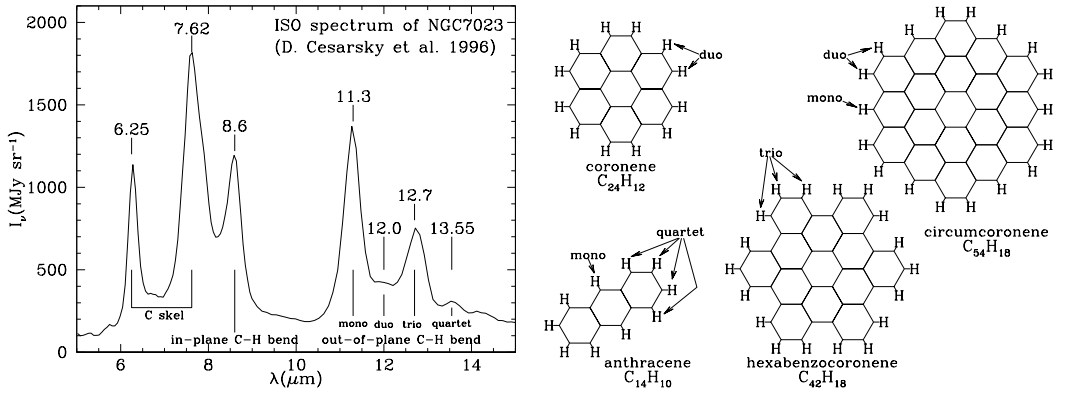


Figure 1.3: *Left panel: The near-IR spectrum of the reflection nebula NGC 7023, with prominent PAH emission features. Adapted from: Cesarsky et al. 1996. Right panel: Structure of four PAHs, with their characteristic aromatic rings. From: Draine*, Physics of the Interstellar and Intergalactic Medium.

PAH molecules turn out to be quite omnipresent and are detected in H II-regions, reflection nebulae (see Fig. 1.3), surfaces of dark clouds, diffuse interstellar clouds and cirrus clouds, surfaces of proto-planetary disks, galactic nuclei, the interstellar medium of galaxies as a whole (see Fig. 1.2), and star burst galaxies[1]. They are very abundant, $\sim 10^{-7}$ relative to hydrogen, locking up to 20 percent of the total cosmic carbon abundance.

PAH molecules are also found in meteorites. When they are seen in meteorites, they almost always have oxygen or deuterium attached to them. These modifications to the PAH molecules probably occur as a result of chemical reactions on the surfaces of grains covered with water ice that are exposed to ultraviolet light. This subclass of PAHs are referred to as quinones, and receive considerable attention by the astrobiology community because they are common to all life forms. They have been suggested to be potentially significant for the formation of life.

PAHs may just be one representative of the molecular universe. In fact, visible spectra of stars usually show prominent absorption features that are too broad to be atomic in origin. These so-called *diffuse interstellar bands* (DIBs) are generally attributed to absorption by moderately large molecules (10–50 C atoms). There are in excess of 400 DIBs known, of which 50 are moderately strong (Hobbs et al. 2009). Because, typically, the visible spectrum of a molecular species is dominated by at most one strong transition, the DIBs implicate the presence of a large number of different molecular species.

[1]PAHs too are very abundant in charbroiled hamburgers and engine soot.

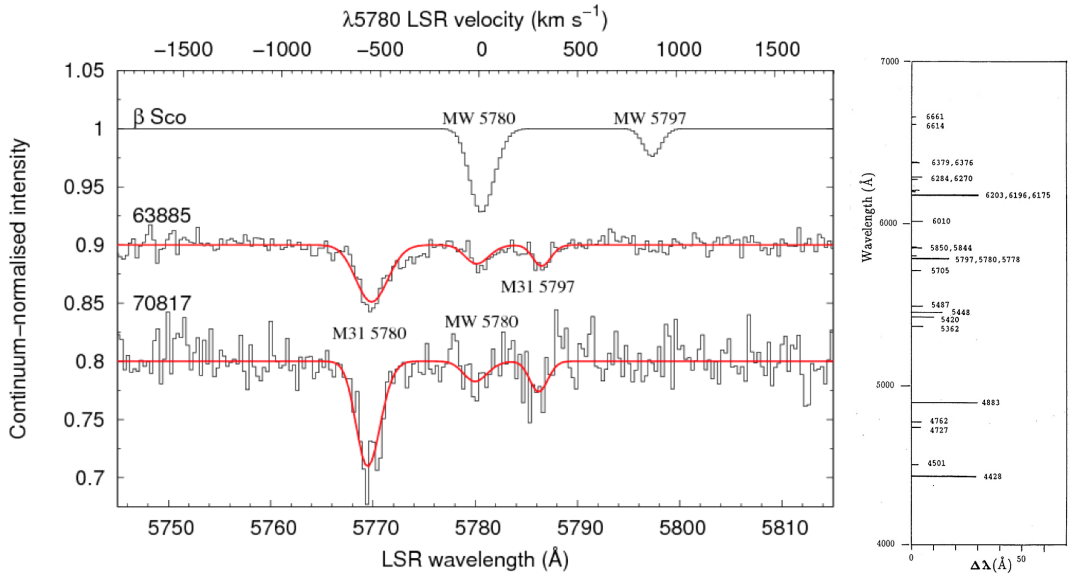


Figure 1.4: *Left panel: The λ5780 Å and λ5795 Diffuse Interstellar Bands (DIBs) in the line of sight toward the galactic (MW) star β Sco and toward stars 63885 and 70817 in M31. Note that towards 70817, also a galactic λ5780 can be seen. The Local Standard of Rest (LSR) velocity scale has been centered on the minimum of the λ5780 DIB. Notice that the width of DIBs is about the same as that of stellar spectral lines, implying that DIBs may be hard to identify in stellar spectra. Right panel: Distribution of main DIBs and their widths. From: N. Cox (left panel); Hoyle & Wickramasinghe (1990) (right panel).*

1.2 Interstellar dust

The most abundant elements, H and He, were formed just after the Big Bang. All other elements formed later in the interiors of stars through nuclear fusion. Interstellar gas is continuously enriched with these heavier elements, through stellar winds and supernovae. We know from observations of interstellar absorption lines in spectra of stars near the Sun that the abundances of the elements in interstellar gas are (much) lower that those in the Sun. This is called *interstellar depletion*. In Fig.1.5 we show the depletion as a function of the condensation temperature of the element, which is the temperature at which that element will 'freeze out' from the gas phase to the solid phase upon cooling of the gas. This process of condensation of solid material, also called dust formation, happens mostly in the cool stellar winds of red giant stars and asymptotic giant branch stars, but also in supernova explosions. The dust that forms in this way contains predominantly metals and very little hydrogen and helium. The figure shows that condensation is in general more complete when it occurs at high temperature. Elements such as titanium, sodium en calcium almost completely freeze out onto grains.

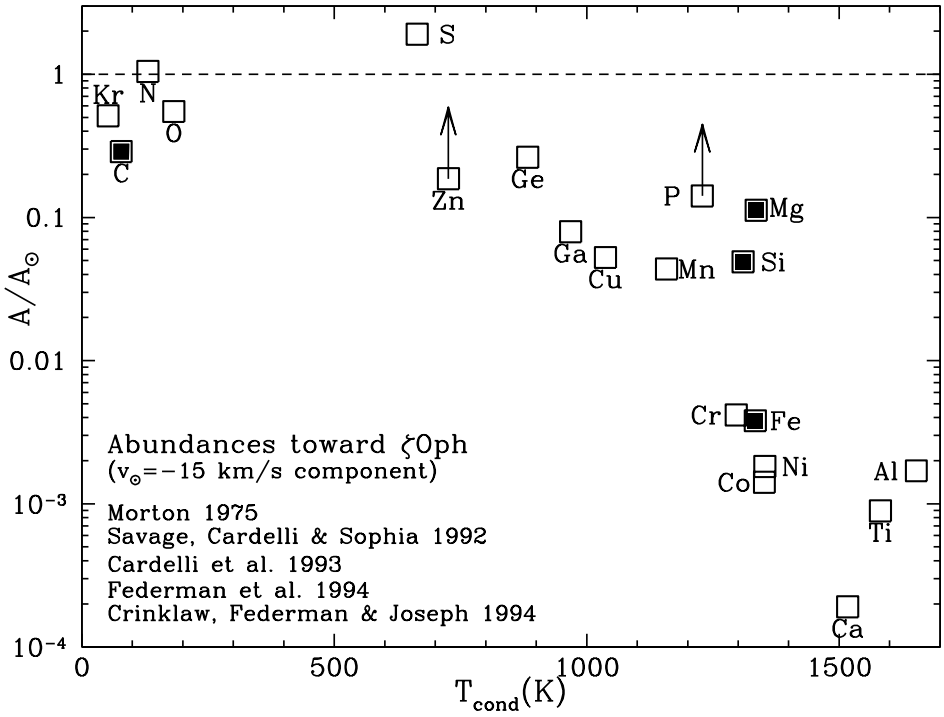


Figure 1.5: *Gas-phase abundances (relative to solar) in the diffuse cloud toward ζ Oph, plotted versus condensation temperature of the element. Solid symbols: major grain constituents C, Mg, Si, Fe. The apparent overabundance of S may be due to observational error, but may also arise because of S II absorption in the H II region around ζ Oph, skewing the solid sulpher abundance determination. The depletion of the high temperature condensates reflects interstellar depletion. From: Bruce Draine,* Physics of the interstellar medium and intragalactic medium.

Dust manifests itself in the interstellar medium in various ways. Through their absorption and scattering properties, small dust grains give rise to a general reddening and extinction of the light from distant stars. Moreover, elongated large dust grains aligned in the galactic magnetic field cause polarization of starlight. Near bright stars, scattering of starlight by dust produces reflection nebula. Finally, the interstellar medium is bright in the infrared because of continuum emission by dust grains.

Interstellar grains are known to contain several populations of very small or nano particles. Nano diamonds have been isolated from meteorites with an isotopic composition that indicates a presolar origin. Likewise, silicon nano particles may be the carrier of a widespread luminescence phenomena, the so-called *extended red emission* (ERE)[2]. In a sense, large PAH

[2]Interstellar dust in nebulae and in the diffuse ISM of galaxies contains a component which responds to illumination by ultraviolet photons with efficient luminescence in the 500 nm to 1000 nm spectral range, known as ERE. No completely satisfactory model for the ERE carrier / process exists at this time. See e.g. Witt & Vijh (2003).

Component		u (eV cm^{-3})
Cosmic microwave background	$T_{\rm CMB}$ = 2.725 K	0.265[a]
Infrared radiation from dust		0.31[b]
Starlight	$h\nu < 13.6$ eV	0.54[c]
Thermal kinetic energy	$(3/2)\, n\, kT$	0.49[d]
Turbulent kinetic energy	$(1/2)\, \rho\, v^2$	0.22[e]
Magnetic energy	$B^2/8\pi$	0.89[f]
Cosmic rays		1.39[g]

Table 1.3: *Sources contributing to the mean total (i.e. frequency integrated) energy density in the solar neighborhood.* [a] *Fixsen & Mather (2002);* [b] *and* [c]*, see chapter 6;* [d] *For* $nT = 3800\,cm^{-3}\,K$*;* [e] *for* $n_{\rm H} = 30\ cm^{-3}$, $v = 1\,{\rm km\,s^{-1}}$, *or* $< n_{\rm H} >= 1\ cm^{-3}$, $< v^2 >^{1/2} = 5.5\,{\rm km\,s^{-1}}$*;* [f] *for a median* $B_{\rm tot} \sim 6.0\,\mu G$ *(Heiles & Crutcher 2005);* [g] *Draine (2011).*

molecules seem to represent the extension of the nano grain size distribution into the molecular domain.

1.3 Energy sources

Energy is present in the ISM in a number of forms: thermal energy, bulk kinetic energy, cosmic ray energy, magnetic energy, and energy in photons. Sources for the latter (see also below) are the cosmic microwave background, far-IR emission from dust, starlight, and emission by hot ionized gas. Table 1.3 provides an overview of the main sources contributing to the energy density in the solar neighborhood. Beware that large fluctuations in for instance the contribution of starlight may occur as a function of location in the galaxy.

Radiation fields

The ISM is permeated by various photon fields, which influence the physical and chemical state of the gas and dust. Figure 1.6 provides an overview of the most important components of the interstellar radiation field in the solar neighborhood. We briefly discuss these component, from the most energetic to the least energetic. In doing so, we closely follow Tielens (2005).

At the shortest wavelengths, emission by hot plasma's – the coronal gas in the halo of our galaxy and in supernova remnants (SNRs) – dominate the radiation field. Numerous emission lines contribute to this component. There is also an extra-galactic contribution at the hardest energies. These X-ray components suffer from absorption by foreground gas, an effect that is least important at the shortest wavelengths. At wavelengths longer than 91.2 nm (or 912 Å), hydrogen is not able to cause continuum absorption through photo-ionization from the

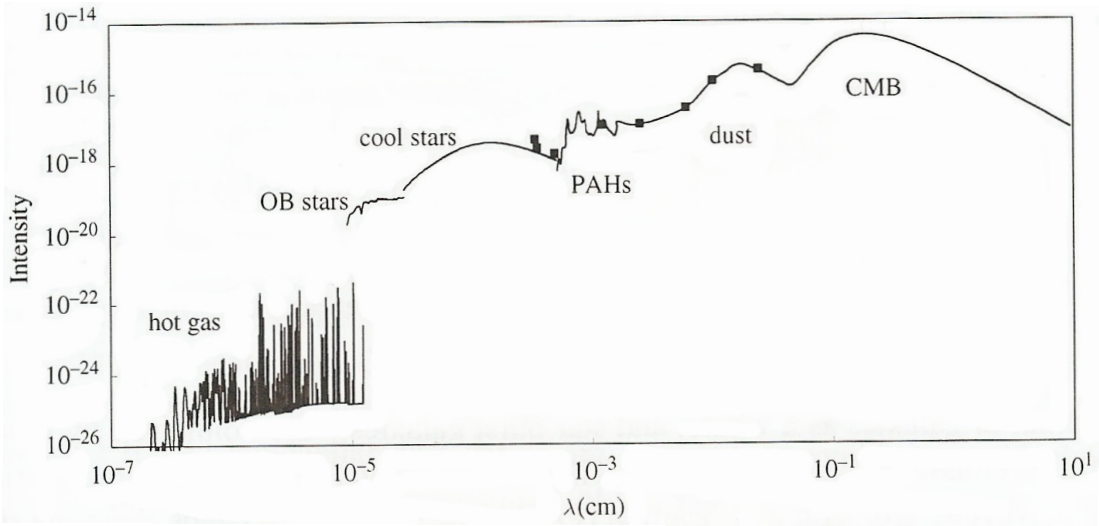


Figure 1.6: *The mean intensity in units of ergs cm^{-3} s^{-1} Hz^{-1} sr^{-1} of the interstellar radiation field in the solar neighborhood. The Lyman continuum edge is at* 9.12×10^{-6} *cm or 91.2 nm. From: Xander Tielens,* Physics and chemistry of the interstellar medium.

ground level, and the intensity of the IS radiation field increases by some 5 to 6 orders of magnitude. Stars now dominate this field. The stellar radiation field contains contributions from O- and B-type stars, which dominate the far-UV wavelengths, A-, F- and G-type stars, which control the visible region, and late-type stars, which are important at far-red to near-IR wavelengths. These stellar photons are absorbed by dust grains that re-radiate this energy at longer wavelengths – in discrete emission bands in the mid-IR and in continuum emission in the far-IR and sub-millimeter regions. The 2.725 K cosmological background dominates at millimeter wavelengths.

Magnetic fields

The interstellar magnetic field is an important source of energy and pressure in the ISM. It consists of a uniform component and a non-uniform component. The uniform magnetic field in the galactic disk generally follows the spiral structure, but with reversals of magnetic field direction from spiral arms to inter-arm regions. The spiral arm magnetic fields are all counterclockwise (as viewed from the North Galactic pole), while the inter-arm fields are clockwise. The spiral arm magnetic field strength appears to be greater than the inter-arm value by a factor of $\sim$1.5. At the location of the sun, at 8.5 kpc from the galactic center, the uniform magnetic field is about 2 μG. At 2 kpc from the galactic center it has increased to about 4 μG. There is also a considerable non-uniform magnetic field, partly associated with expanding interstellar shells (super-bubbles) and their shocks. The total magnetic field in the solar neighborhood is $\sim$5 μG.

Inside dense clouds the strength of the magnetic field B increases, roughly with the square root of the particle density. At $n \simeq 10^4$ cm^{-3} a typical value is $B \simeq 30\,\mu$G.

Cosmic rays

Cosmic rays (CRs) are very high-energy particles, that almost always originate from outside of the solar system. The term 'ray' is a historical accident, as cosmic rays were at first mistakenly thought to be mostly electromagnetic radiation. Cosmic rays are tied to the galactic magnetic field and confined to a disk of radius $\sim$12 kpc with a thickness of $\sim$2 kpc.

Particles with energies $\gtrsim 100$ MeV per nucleon contribute considerably to the energy density of the ISM. Low-energy ($\simeq 100$ MeV) cosmic rays are important for the heating and ionization of interstellar gas. Unfortunately, it is difficult to measure the flux of low-energy CRs within the heliosphere because of strong modulation by the solar wind. These low-energy cosmic rays may ionize atoms in the interstellar medium, contributing in an important way to the formation of simple molecules (such as OH).

Cosmic rays consist mainly of relativistic protons and α-particles with energies in the range of 1–10 GeV. The abundance pattern of the elements in the cosmic rays is non-solar. A wide variety of sources may be responsible for the production of these relativistic particles, including supernovae, active galactic nuclei, quasars, and gamma-ray bursts. Part of the cosmic rays may interact with interstellar matter to form other elements. For instance, carbon, nitrogen and oxygen nuclei may collide with interstellar matter – gas, dust, or other cosmic rays – to form lithium, beryllium and boron. This process is referred to as cosmic-ray spallation. Interaction of the cosmic-ray protons with interstellar gas may also produce gamma rays with $E_\gamma \gtrsim 50$ MeV through π^0 meson decay emission. Likewise, the interaction of energetic ($<$ 1 GeV) electrons with interstellar gas gives rise to gamma rays through bremsstrahlung and inverse Compton scattering.

Kinetic energy of the ISM

Stellar outflows and supernovae explosions supply kinetic energy to the ISM. The expanding shells blown by individual stars and the superbubbles blown by the concerted action of OB associations shape the ISM. They sweep up and compress the ambient ISM and set it into motion. These motions are often Rayleigh-Taylor and Kelvin-Helmholtz unstable and, in general, are converted into turbulence. When clouds collide, shocks may be produced that heat the gas. This gas may in turn cool through line radiation. Though important for the energy density budget of the ISM, the total mechanical energy output of stars (outflows and supernovae) is only small ($\sim$ 0.5 percent) compared to the total stellar radiative energy output.

Turbulence in individual molecular clouds (so, we're now looking at small scales, of fractions of a parsec) is probably of a magneto-hydrodynamic nature. It helps to support these clouds against self-gravity. This turbulent energy is tapped from the magnetic or rotational energy

supporting the cloud or is supplied by powerful outflows driven by newly formed stars in the nearby surroundings.

Equipartition of energy densities

It seems quite remarkable that the energy densities of the energy sources discussed above all fall within the range $0.2 - 2\,\mathrm{eV\,cm^{-3}}$ (see Table 1.3). Is this near-equipartition coincidental? The fact that the energy density in the CMB is similar to the other energy densities is surely accidental. The other energy densities likely are coupled. In this discussion we follow Draine (2011).

The magnetic energy has been built up by fluid motions, so it is probably not a coincidence that the magnetic energy density $B^2/8\pi$ and the turbulent energy density $(1/2)\rho v^2$ are comparable in magnitude. Similarly, if the cosmic-ray energy were much larger, it would not be possible for the magnetized ISM to confine the cosmic rays, and they would be able to escape freely from the Galaxy. This limits the cosmic-ray energy density to approximately equipartition with the sum of the turbulent energy density and thermal energy density in the ISM.

That the energy in starlight is comparable to the gas pressure may be coincidental. However, if the starlight energy density were much larger (by a factor $\sim 10^2$), radiation pressure acting on dust grains would be able to 'levitate' the ISM above and below the galactic mid-plane, presumably suppressing star formation. This 'feedback loop' may play a role in regulating the starlight energy density in star-forming galaxies.

The energy sources discussed so far heat the gas and dust. That the energy densities of these (thermal kinetic energy and infrared radiation, respectively) are comparable to that of their heating sources implies that this heating is quite efficient.

1.4 Galactic chemical evolution – gas and dust ejection into the ISM

The accretion of primordial gas from a reservoir in a halo is assumed to form a galaxy. The first galaxies likely started their formation several hundred million years after the Big Bang, which produced primordial baryonic matter in the form of hydrogen (H) and helium (He) in the first 15 minutes after the start of universe formation. Let us discuss the conceptual outline of a galaxy formation model[3] to get a rough idea of the physics involved. This physics describes the formation of stars, the production of elements in their stellar centers en the partial return of primordial and new elements to the interstellar medium (a process called nucleosynthetic feedback) causing chemical enrichment of the interstellar medium of the galaxy. Then, let us skip to the results of such a model to assess, over cosmic time, the production rate and origin of the elements heavier then helium present in the interstellar medium of the galaxy.

[3]For more details, see Kobayashi et al. 2000, ApJ 539, 26.

We assume the halo to contain a baryonic mass $M_{\rm tot}$. The rate of infall of the primordial halo gas to what is to become the galaxy is $R_{\rm in}$ (dimension [s^{-1}]), for which often an exponential form is assumed

$$R_{\rm in} = \frac{1}{\tau_{\rm in}} e^{-t/\tau_{\rm in}}. \tag{1.1}$$

Here $\tau_{\rm in}$ is a characteristic infall timescale. So, if $f_{\rm bd}$ is the ratio of the gas mass in the galactic bulge and disk to the total baryon mass of the galaxy (bulge, disk and halo), it holds that

$$\frac{{\rm d}f_{\rm bd}}{{\rm d}t} = R_{\rm in}. \tag{1.2}$$

The star-formation rate ψ is assumed to be proportional to the gas fraction as

$$\psi = \frac{1}{\tau_{\rm s}} f_{\rm bd}, \tag{1.3}$$

where $\tau_{\rm s}$ is the star formation timescale. For (what is to become) elliptical galaxies this time scale is likely quite short (~ 0.1 Gyr); for spiral galaxies it is longer (~ 2 Gyr). The gas mass that is turned into stars reduces the gas mass of the bulge and disk. Feedback processes again add (nucleosynthetically enriched) material to that gas mass. This results in

$$\frac{{\rm d}f_{\rm bd}}{{\rm d}t} = -\psi + E_{\rm high} + E_{\rm low} + R_{\rm in}, \tag{1.4}$$

where $E_{\rm high}$ is the rate at which gas is ejected into the interstellar medium by mass loss and type-II core-collapse supernovae from massive stars, and $E_{\rm low}$ is the rate at which gas is ejected into the interstellar medium by mass loss and type Ia supernovae from low-mass stars. Using input from stellar evolution models and supernova nucleosynthesis models one may specify how each of these processes contributes to the mass fraction Z_i of heavy element i in the gas, such that one may derive the time variation of Z_i in the galactic gas, i.e.,

$$\frac{{\rm d}(Z_i \cdot f_{\rm bd})}{{\rm d}t} = -Z_i \cdot \psi + E_{\rm high}(Z_i) + E_{\rm low}(Z_i) + Z_{i,{\rm in}} \cdot R_{\rm in}, \tag{1.5}$$

where the metallicity of the infall gas $Z_{i,{\rm in}}$ is assumed to be zero when considering galaxy formation at early cosmic times. The terms $E_{\rm high}$ and $E_{\rm low}$ depend on the initial mass function (of single stars and primaries in binary systems), the mass ratio and period distribution in binaries, and the SN Ia rate[4], which is controlled by binary evolution. An additional term may be introduced to describe element production and feedback in compact object mergers, such as neutron star - neutron star mergers.

Figure 1.7 shows a model of the cosmic production of elements in the Milky-Way galaxy. The results should be taken with care as much of the physics still is uncertain. Main uncertainties are in the specifics of binary evolution, mass loss (through stellar winds, eruptive

[4]SN Ia occur in binary systems in which one component is a white dwarf. By accreting gas from its companion (a main sequence star, giant, or more massive white dwarf) it may reach a critical mass (the Chandrasekhar mass) that initiates a supernova explosion.

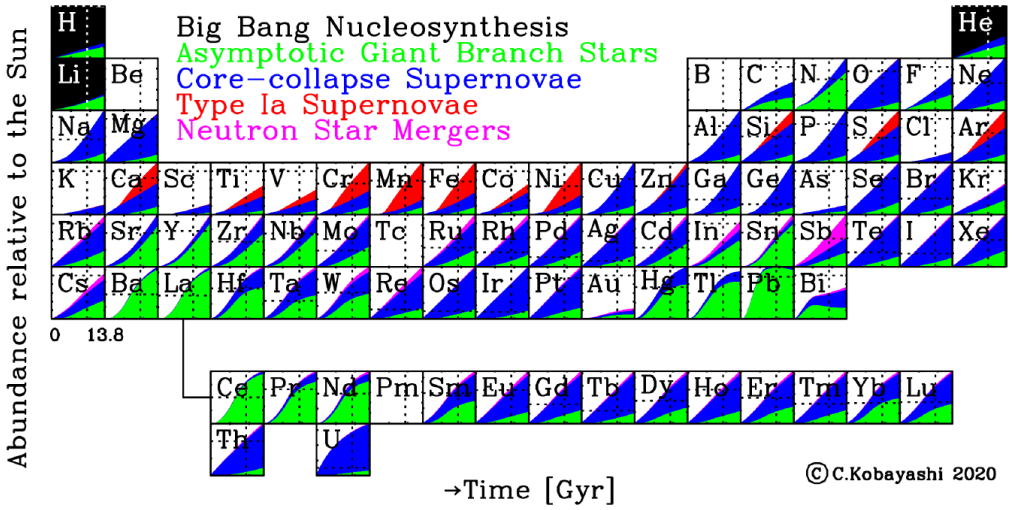


Figure 1.7: *The cosmic production of the elements from Galactic chemical evolution models by Kobayashi et al. (2020, ApJ 900, 179). In the background of each element box in the periode table the main producers of the element are shown as a function of cosmic time (see legend). The amounts returned via mass loss are also included for Asymptotic Giant Branch (AGB) stars and stars ending as core-collapse supernovae. The x-axis of each box shows time from $t = 0$ (Big Bang) to 13.9 Gyrs, while the y-axis shows the linear abundance relative to the sun, $X/X_\odot$. The dotted lines indicate the observed solar value, i.e., $X/X_\odot = 1$ (horizontal dotted line) and 4.6 Gyrs (vertical dotted line) for the age of the sun (measured back from the present time). For H, He, and Li, the destruction of the element is shown by either AGB or massive stars.*

events, binary interaction, and SN explosions), and nucleosynthetic reaction rates (e.g., the $^{12}\mathrm{C}(\alpha,\gamma)^{16}\mathrm{O}$ rate). The contribution of low mass stars (indicated in green) to the chemical enrichment occurs almost exclusively during the Asymptotic Giant Branch (AGB) phase where the star may loose up to $\sim 85\,\%$ of its initial mass. The contribution of massive stars is given in red and includes feedback in stellar winds and in their life ending core-collapse supernovae. A complex set of channels produces the elements up to iron (Fe), including hydrogen burning in the p-p chain and CNO cycle, NeNa and MgAl cycles, the triple-α process, C- and O-fusion, Ne-disintegration, and photo-disintegration and α-capture re-assembly[5]. Elements heavier than Fe are formed through recursive neutron capture processes. If the capture rate of neutrons is 'slow' (passing decades between neutron captures) the rate at which an iron-group seed particle captures neutrons is so low that there is sufficient time for β-decay to occur before another neutron is captured, producing an element of the next higher atomic number. Such processes occur in AGB stars (where the $^{13}\mathrm{C}(\alpha,n)^{16}\mathrm{O}$ reaction is the neutron source) and at the end of He- and C-burning in massive stars (where $^{22}\mathrm{Ne}(\alpha,n)^{25}\mathrm{Mg}$ is the neutron source). In explosive processes the neutron capture rate can be so 'rapid' (with seconds in between) that unstable isotopes that are formed have no time to decay between neutron captures. Captures ensue as long as the neutron flux is high enough, for instance, in the onset

[5]See, e.g., Lamers & Levesque, *Understanding Stellar Evolution*.

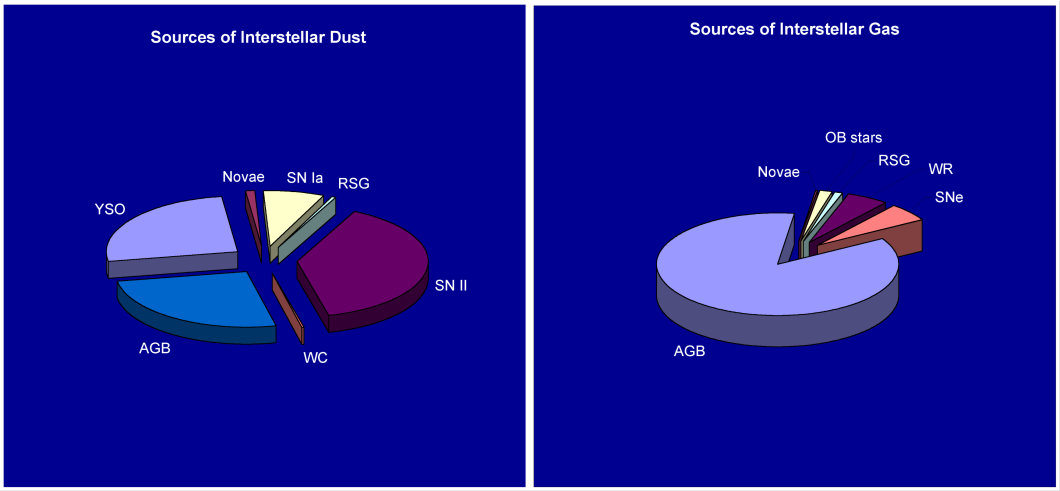


Figure 1.8: *Stellar sources of gas and dust. From: Barlow 2009*, Astrophysics in the Next Decade.

of core-collapse supernova, SN Ia, or neutron star mergers. When the neutron flux stops the neutron-rich isotopes will suffer a series of β-decays until a stable isotope is reached.

Notice that the figure 1.7 does not report on the origin of Beryllium and Boron. The dominant production channel of these elements is cosmic ray spallation, i.e., the collision between an atom in space (often carbon, nitrogen and oxygen) and a relativistic particle, a process not considered by Kobayashi and co-authors.

Table 1.4 summarizes current gas and dust mass injection rates into the ISM of our own Milky Way for a variety of stellar objects, following Tielens (2005). The dust budget has been split out into two separate columns according to whether the surface layers of the stellar source are more abundant in carbon than in oxygen (C/O $>$ 1), which leads to carbonaceous dust formation, or more abundant in oxygen than in carbon (C/O $<$ 1), which leads to the formation of oxides, notably silicates, or metals. Again, the quoted numbers are uncertain; this remains a vivid field of research, and new developments have been known to change our ideas considerably.

The relative importance of the gas and dust sources listed in Table 1.4 vary across the Galaxy, for instance because of the general increase in metallicity towards the galactic center. This causes the ratio of O-rich to C-rich (asymptotic) giants to increase towards the inner galaxy, as well as the Wolf-Rayet star population.

The gas mass return rate is expressed in solar masses per million years per square kpc, so the rate is integrated over the thickness of the galactic disk. It is dominated by evolved low-mass (asymptotic) cool giants, as expected, since low-mass stars dominate the stellar mass of the Galaxy. On the AGB low-mass stars inject much carbon formed by the triple-α process. Type Ia supernovae, which also have low mass progenitors, inject considerable amounts of iron into

Source	$\dot{M}$(gas)	$\dot{M}$(C dust)	$\dot{M}$(Silicate & metal dust)
		$M_\odot$ kpc^{-2} Myear^{-1}	
C-rich (asymptotic) giants	750	3.0	–
O-rich (asymptotic) gians	750	–	5.0
Novae	6	0.3	0.03
SN type Ia	–	0.3	2
OB stars	30	–	–
Red Supergiants	20	–	0.2
Wolf-Rayet stars	100	0.06	–
SN type II	100	2	10
Star formation	-3000	–	–
Disk-Halo circulation	7000		
Infall from satellite galaxies	150		

Table 1.4: *Injection and depletion rates of interstellar gas (2nd column) and dust (3rd and 4th column). From: Tielens,* The physics and chemistry of the interstellar medium.

the ISM. Massive stars are more efficient in producing and injecting elements such as oxygen and silicon, into the ISM.

The disk of the Galaxy also exchanges material with the lower halo. The concerted action of supernovae set up a galactic fountain, removing material from the galactic disk. Material from the lower halo falls back to the galactic disk at about the same rate. At 7000 $M_\odot$ kpc^{-2} Myear^{-1} this is about 5 $M_\odot$ of material per year adopting a radius of the galactic disk of 15 kpc (such that the galactic disk surface is 700 kpc^2). In terms of the mass flux, this circulation pattern dominates the mass budget.

The Magellanic Stream, a gas stream from the Magellanic Clouds to our Milky Way as a result of tidal effects, supplies some 150 $M_\odot$ kpc^{-2} Myear^{-1}.

There are also some indirect arguments that suggest that the Galaxy may have been accreting about $0.5 M_\odot \mathrm{yr}^{-1}$ of metal-poor primordial gas over much of its lifetime (see Fig. 1.1) or about an order of magnitude more than the Magellanic Stream accretion. This last component has not been taken into account in Table 1.4.

1.5 The lifecycle of the galaxy

The origin and evolution of galaxies are closely tied to the cyclic processes in which stars form from the gas and dust reservoir that is the interstellar medium, and return gas and dust to this entity through outflows and explosions. We follow Tielens (2005) in this short discussion. Winds from low-mass stars – the number of which is a function of the past star formation rate – control the total mass balance of interstellar gas and contribute substantially to the injection

of dust, an important opacity source, and PAHs, an important heating agent of interstellar gas. Winds and supernovae explosions from high-mass stars – the number of which is a function of the present star formation rate – dominate the mechanical energy injection into the ISM, and thus the turbulent pressure that helps support clouds against galactic- and self-gravity. Through the formation of the hot coronal phase (see section 1.1), massive stars regulate the thermal pressure as well.

Massive stars also control the far-UV photon energy budget, which is an important heating, ionizing, and dissociation source of the interstellar gas, and they are also an important source of intermediate-mass elements (specifically oxygen and silicon) that are an important component of interstellar dust. Eventually, it is the dust opacity – the ability of the grains to shield the medium against UV photons – that allows molecule formation and survival. The cooling capability of molecules is crucial in the onset of gravitational instability of molecular clouds.

In the star formation process, interstellar dust ends up in the disk around the forming stars. These grains are the building blocks of planets, and potentially life.

Clearly, therefore, there is a complex system of feedback between star and planet formation and the ISM. And it is this feedback that determines the structure, composition, chemical evolution, and observational characteristics of the interstellar medium in the Milky Way and in other galaxies all the way back to the first stars and galaxies that formed at redshifts $z \gtrsim 10$.

If we want to understand this interaction, we have to understand the fundamental physical processes that govern the ISM. These processes are the topic of these lectures.

1.6 Literature

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

- Spitzer (Lyman Jr.) *Physical processes in the interstellar medium*, 1998 Edition. Wiley. Outstanding book discussing the physical processes in the interstellar medium and the nature of interstellar matter, with a strong emphasis on basic physical principles.
- Waters (Rens), with contributions from Kama (Mihkel) *Lecture notes of the master course: Interstellar and Circumstellar Matter*, 2013 version. Master course lecture notes of the University of Amsterdam. Very clear and conceptual. Borrows from, and used as a main guide to write several chapters of these lecture notes
- Lequeux (James): *The Interstellar Medium*, 2003. Springer. Excellent textbook with in depth discussions of the derived formulae. Clear, consice and easy to read.
- Draine (Bruce T.): *Physics of the interstellar and intergalactic mediam*, 1st Edition 2011. Princeton University Press. Excellent new and accessible book with lots of useful

data. Borrowed from, and used as a main guide to write several chapters of these lecture notes.

- Tielens (Xander G.G.M.): *The physics and chemistry of the interstellar medium*, 1st paperback edition 2010. Princeton University Press. Excellent book focussing on the detailed physics of gas and dust, with special attention for polycyclic aromatic hydrocarbons.

- Dyson (J. E.) and Williams (D. A.): *The physics of the interstellar medium*, 2nd revised edition 1997 Taylor & Francis Ltd. Excellent book aiming at advanced undergraduates.

- Osterbrock (Donald E.) and Gerland (Gary J.): *Astrophysics of gaseous nebulae and active galactic nuclei*, 2nd edition 2006 University Science Books. Standard textbook on the physics of nebulae and AGNs. Beautifully written.

- Stahler (Steven W.) and Palla (Francesco): *The Formation of Stars*, 1st Edition 2004 (2nd Reprint 2011). Wiley-VCH Verlag GmbH & Co, KGaA. Excellent standard book, focussing on star formation.

- Evans (Aneurin): *The Dusty Universe*, 1994 John Wiley & Sons Ltd. Introduction on dust in the universe, written at an advanced undergraduate level, but very worthy of reading.

- Bransden (B. H.) and Joachain (C. J.): *Physics of atoms and molecules*, 2nd Edition 2003 New York, Prentice Hall. This book deals with both atomic and molecular physics, at an advanced undergraduate level.

Exercise 1.1

The ten most important elements in life on Earth are H, C, N, O, Na, P, S, Cl, K, and Cl (see *An introduction to Astrobiology*, Revised edition 2011, Cambridge University Press). By percent mass they contribute about 9.5, 18.5, 3.2, 65, 0.2, 1.0, 0.3, 0.2, 0.4, and 1.5 %, respectively, to the human body.

a) Which percentage do these building blocks of life, save H, represent of the total of elements heavier than helium? (Use an abundance table from the literature). Are stars 'geared' towards producing the elements necessary for life on Earth or are these elements only a byproduct of stellar nucleosynthesis?

b) Figure 1.7 shows the time evolution (in Gyr) of the origin of the elements in the periodic table. Estimate (roughly) which percentage of the human body (by mass) originates from Big Bang nucleosynthesis, AGB stars, core-collapse Supernovae (so, from massive stars) and Type Ia supernovae (so, from low-mass stars).

2

Radiation and matter

In this chapter we review the basic quantities describing the radiation field and aspects of radiation transfer. The description of the coupling between the gas and dust and the radiation field can be simplified considerably if the medium and the radiation field are in some type of equilibrium. The most strict form of equilibrium is *thermodynamic equilibrium* (TE). We discuss the properties of a TE medium, and briefly introduce local-TE (LTE) and non-local TE (NLTE).

2.1 Characterizing radiation fields

In this section we give the basic definitions that characterize a non-polarized radiation field.

Specific intensity

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

$$\begin{aligned} dE_\nu &= I_\nu(\mathbf{r}, \mathbf{n}, t)\ \mathbf{n} \cdot \mathbf{dS}\ d\omega d\nu dt \\ &= I_\nu(\mathbf{r}, \mathbf{n}, t)\ \cos\theta\ dS\ d\omega d\nu dt \end{aligned} \tag{2.1}$$

where θ is the angle between the direction of the beam and the normal to the surface, i.e. $\mathbf{n} \cdot \mathbf{dS} = \mathbf{n} \cdot \mathbf{s} dS = \cos\theta\ dS$ (see figure 2.1). For a beam entering the medium at a polar angle θ between the normal direction and the beam direction, we will frequently use the variable $\mu = \cos\theta$. The dimensions of I_ν are erg cm^{-2} sec^{-1} hz^{-1} sr^{-1}. From a macroscopic point of view, the specific intensity provides a complete description of the (non-polarized) radiation field.

The specific intensity is defined such that it is independent of distance if there are no sources or sinks of radiation along the direction of the beam. This implies that the value for the specific

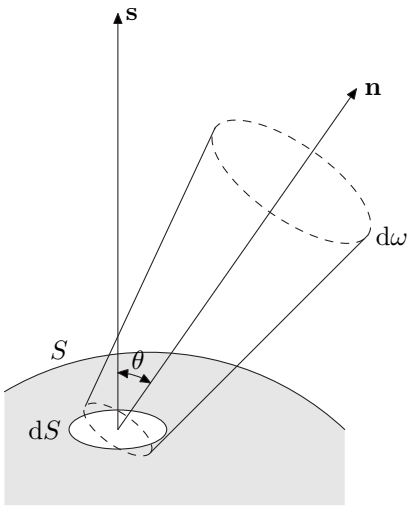


Figure 2.1: *Definition of the specific intensity as a beam of radiation into a solid angel $d\omega$, across a surface dS oriented in direction* $\mathbf{s}$*. The vector* $\mathbf{n}$ *is the direction of propagation of the beam which is at an angle θ with* $\mathbf{s}$*.*

intensity at the source location can be obtained by measuring the amount of energy per unit time and per frequency interval that hits a detector (with known efficiency) per unit detector surface, when the solid angle subtended by the source is known. To be able to measure the specific intensity it is therefore required that the source is spatially resolved.

The invariance of I_ν

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

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

We can understand this property by considering a beam of radiation that passes through a surface element dS at position $\mathbf{r}$, as well as through an element dS' at position $\mathbf{r}'$ (see figure 2.2). The amount of energy passing through both areas is

$$dE_\nu = I_\nu \cos\theta \, dS \, d\omega d\nu dt = I'_\nu \cos\theta' \, dS' \, d\omega' d\nu dt \tag{2.2}$$

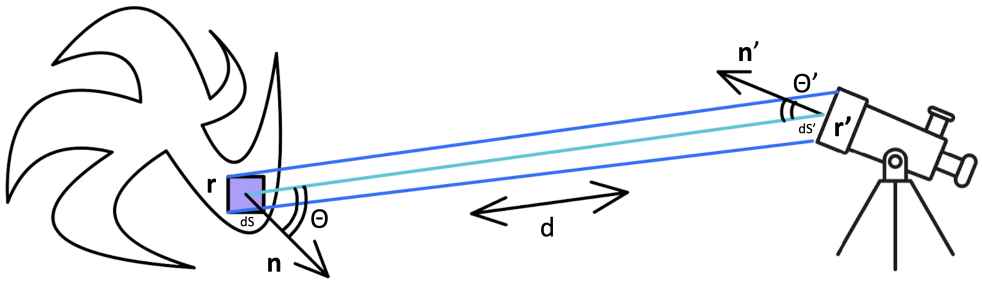


Figure 2.2: *Illustration of the proof of invariance of the specific intensity. The points* $\mathbf{r}$ *and* $\mathbf{r}'$ *are separated by a distance* d*. Area* dS *subtends a solid angle* $d\omega'$ *seen from position* $\mathbf{r}'$*, and the area* dS' *subtends* $d\omega$ *at* $\mathbf{r}$*.* $\mathbf{n}$ *and* $\mathbf{n}'$ *are unit vectors normal to* dS *and* dS'*.*

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

Mean intensity

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

$$J_\nu(\mathbf{r}, t) = \frac{1}{4\pi} \oint I_\nu(\mathbf{r}, \mathbf{n}, t)\, d\omega \tag{2.3}$$

The mean intensity is the zero-order moment of the specific intensity and has dimensions erg cm^{-2} sec^{-1} hz^{-1}. In spherical coordinates $d\omega = \sin\theta d\theta d\phi = -d\mu d\phi$. (see figure 2.3). The total solid angle Ω therefore is

$$\Omega = \oint d\omega = \int_0^{2\pi} \int_0^{\pi} \sin\theta\ d\theta d\phi = 2\pi \int_{-1}^{+1} d\mu = 4\pi \tag{2.4}$$

This explains the normalization factor 4π in eq. (2.3). Solid angle is expressed in the dimensionless unit steradian (sr). In the small angle limit, i.e. $\theta_\star \to 0$ or $\mu_\star = \cos\theta_\star \to 1$ of the axially symmetric case (for instance, we observe a circular looking object in the sky), we find

$$\Omega_\star = 2\pi \int_0^{\theta_\star} \sin\theta\ d\theta = 2\pi\,(1 - \cos\theta_\star) \simeq 2\pi\left(1 - \left[1 - \theta_\star^2/2\right]\right) = \pi\,\theta_\star^2, \tag{2.5}$$

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

Notice that though the mean intensity is defined per steradian, through the factor $1/4\pi$ in eq. 2.3, its dimension does not convey this. The reasoning behind this is that J_ν is no longer a function of solid angle; as solid angle is a dimensionless unit explicit reference to it may be dropped.

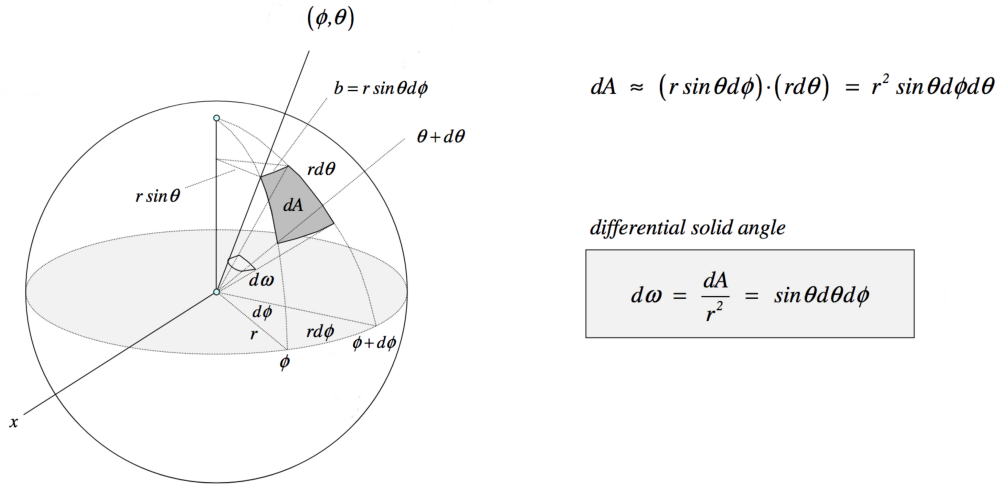


Figure 2.3: *The infinitesimal solid angle $d\omega$ expressed in spherical coordinates. The area A of a sphere of radius r in the interval $(\theta, \theta + d\theta)$ and $(\phi, \phi + d\phi)$ is $A = b \cdot a = r^2 \sin\theta d\theta d\phi = r^2 d\omega$, so that $d\omega = \sin\theta d\theta d\phi$. The total solid angle is 4π steradians.*

We may now rewrite eq. (2.3) as

$$\begin{aligned} J_\nu &= \frac{1}{4\pi} \int_0^{2\pi} \int_0^{\pi} I_\nu \sin\theta \, d\theta d\phi \\ &= \frac{1}{2} \int_{-1}^{+1} I_\nu \, d\mu \end{aligned} \tag{2.6}$$

The mean intensity is e.g. used in the description of processes such as photoexcitation and photoionization, which only depend on the number of photons at some position at some time and do not depend on the direction of origin of these photons. For an isotropic radiation field $J_\nu = I_\nu$.

Geometrical dilution

Consider a spherical star that emits an isotropic radiation field $I_\nu(\theta, \phi) = I_\nu$ from its surface at $R_\star$. We are interested in the mean intensity above the stellar surface, i.e. at $r > R_\star$. Figure 2.4 shows the directions from which the point r receives the stellar intensity I_ν. For the mean intensity in this point we find

$$J_\nu(r) = \frac{1}{2} \int_{\mu_\star}^{1} I_\nu \, d\mu = \frac{1}{2}(1 - \mu_\star) \, I_\nu \equiv W(r) \, I_\nu \tag{2.7}$$

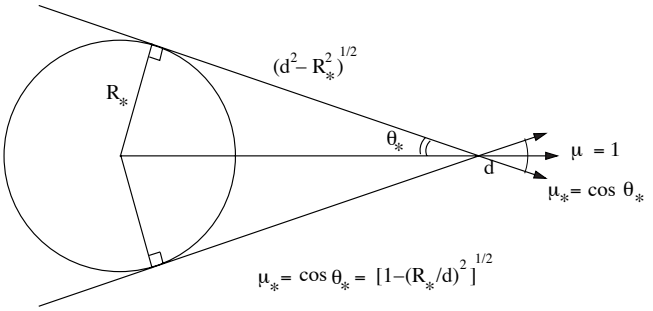


Figure 2.4: *Geometry showing the directions from which a point d receives the stellar intensity I_ν.*

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

$$W(r) = \frac{1}{2}\left\{1 - \left[1 - \left(\frac{R_*}{r}\right)^2\right]^{1/2}\right\} \tag{2.8}$$

The factor W is called *geometrical dilution* and denotes the fraction of the total solid angle subtended by a star of radius $R_\star$ seen from a point in the sky at distance r. If $r = R_\star$ one finds $W = 1/2$. This is easy to understand as an observer on the stellar surface sees half of the sky filled by the star. At very large distances, i.e. $r \gg R_\star$, one may approximate the dilution factor by

$$W(r) \simeq \frac{1}{4}\left(\frac{R_\star}{r}\right)^2. \tag{2.9}$$

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

Energy density of radiation

The energy density of radiation is the amount of energy present in a region of space per unit volume. Let us consider the energy in a beam of solid angle $d\omega$ passing through a surface element dS, i.e., eq. (2.1). We envision a volume V such that for our beam the path length to the far edge of the volume is l. We are only interested in those photons in the beam that are actually inside V. These will be inside of the elementary volume during a time $dt = l/c$. The part of the volume in which the photons are located is $dV = l\, dS \cos\theta$, where θ is the angle between the direction of the beam and the normal to the surface element dS. The energy in this part of the volume provided by the beam is therefore $dE_\nu = c^{-1}\, I_\nu(\mathbf{r}, \mathbf{n}, t)\, d\omega\, d\nu\, dV$. Integrating over all beams (to cover the entire volume) and the volume (to account for the size of the volume through the path length l), one finds for the total energy in V in the frequency band $d\nu$

$$E_\nu(\mathbf{r}, t)\, d\nu = \frac{1}{c}\left[\int_V \left\{\oint I_\nu(\mathbf{r}, \mathbf{n}, t)\, d\omega\right\} dV\right] d\nu \tag{2.10}$$

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

$$u_\nu(\mathbf{r},t) = \frac{E_\nu(\mathbf{r},t)}{V} = \frac{1}{c}\oint I_\nu(\mathbf{r},\mathbf{n},t)\,d\omega = \frac{4\pi}{c}J_\nu(\mathbf{r},t) \tag{2.11}$$

The frequency integrated or *total energy density* is

$$u = \int_0^\infty u_\nu\,d\nu = \frac{4\pi}{c}\,J. \tag{2.12}$$

The dimension of u_ν is erg cm^{-3} Hz^{-1}; that of u is erg cm^{-3}. In studies of the interstellar medium is it quite customary to express the total energy density in eV cm^{-3} (see for instance Table 1.3). Recall that 1 erg cm^{-3} = 6.415×10^{11} eV cm^{-3}.

Flux

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

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

$$\boldsymbol{\mathcal{F}}_\nu(\mathbf{r},t)\cdot\mathbf{dS}\,d\nu\,dt = \oint dE_\nu = \oint I_\nu(\mathbf{r},\mathbf{n},t)\,\mathbf{n}\cdot\mathbf{dS}\,d\omega\,d\nu\,dt, \tag{2.13}$$

where the integration is over solid angle, consequently

$$\boldsymbol{\mathcal{F}}_\nu(\mathbf{r},t) = \oint I_\nu(\mathbf{r},\mathbf{n},t)\,\mathbf{n}\,d\omega, \tag{2.14}$$

that is

$$(\mathcal{F}_x,\mathcal{F}_y,\mathcal{F}_z) = \left(\oint I_\nu(\mathbf{r},\mathbf{n},t)\,\mathrm{n}_x\,d\omega, \oint I_\nu(\mathbf{r},\mathbf{n},t)\,\mathrm{n}_y\,d\omega, \oint I_\nu(\mathbf{r},\mathbf{n},t)\,\mathrm{n}_z\,d\omega\right). \tag{2.15}$$

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

In a plan parallel medium only the flux in the z direction $\mathcal{F}_z \neq 0$. Symmetry arguments show that in the x and y directions the flux $\mathcal{F}_x = \mathcal{F}_y = 0$. As in this case only the z component of

the flux is relevant one usually denotes this flux vector component as 'the' flux. It follows that

$$\begin{aligned}\mathcal{F}_\nu(z,t) &= \int_0^{2\pi}\int_0^{\pi} I_\nu \cos\theta \sin\theta \; d\theta d\phi \\ &= 2\pi \int_{-1}^{+1} I_\nu(z,\mu,t)\mu \; d\mu \end{aligned} \tag{2.16}$$

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

The frequency integrated or *total flux* is

$$\mathcal{F} \equiv \int_0^\infty \mathcal{F}_\nu \, d\nu \equiv \sigma T_{\rm eff}^4 \tag{2.17}$$

where the last equality introduces the *effective temperature* $T_{\rm eff}$. Note that conform its definition, $T_{\rm eff}$ is a measure of the total flux; it is not a physical temperature. When considering a star of radius $R_\star$, integration of the flux over all of the stellar surface results in the *monochromatic luminosity* $L_\nu = 4\pi R_\star^2 \mathcal{F}_\nu(R_\star)$. Further integration over all frequencies yields the *luminosity* of the star

$$L_\star = \int_0^\infty L_\nu \, d\nu = 4\pi R_\star^2 \mathcal{F}(R_\star) = 4\pi R_\star^2 \sigma T_{\rm eff}^4 \tag{2.18}$$

The r^{-2} dependence of the flux

The observational meaning of the flux can be understood in the following way: Consider a constant, isotropically radiating source, e.g. a spherical star with radius $R_\star$ and surface $4\pi R_\star^2$. When we place a concentric spherical surface around this source, with radius r – and if we assume there is no absorption or emission in the space around the source – the total amount of radiation energy passing through the outer surface will be the same as that passing through the stellar surface. Therefore

$$\mathcal{F}_\nu(R_\star)\, 4\pi R_\star^2 = \mathcal{F}_\nu(r)\, 4\pi r^2 \tag{2.19}$$

If r is the distance of the source to earth, then $\mathcal{F}_\nu(r)$ is the observed flux. The flux therefore decreases with distance as $\mathcal{F}_\nu(r) \propto r^{-2}$.

At first sight it may seem that this result is in contradiction to the invariance of the specific intensity along the line of sight. This is not so. Again consider a constant, isotropically radiating star with radius $R_\star$ and emerging intensity I_ν. At a point r the intensity is equal to I_ν for all beams that intersect the star, for all other beams it is zero. We then find for the flux in r (see figure 2.4)

$$\mathcal{F}_\nu(r) = 2\pi \int_{\mu_\star}^1 I_\nu \mu \, d\mu = \pi I_\nu (1-\mu_\star^2) = \pi I_\nu \left(\frac{R_\star}{r}\right)^2 \tag{2.20}$$

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

$$\mathcal{F}_\nu(R_\star) = 2\pi \int_{\mu_\star=0}^{1} I_\nu \mu \, d\mu \equiv \mathcal{F}_\nu^+ = \pi I_\nu \tag{2.21}$$

Angular diameter

The *angular diameter* of an object of radius $R_\star$ is

$$\alpha_\star = \frac{2R_\star}{d} \tag{2.22}$$

such that the observed flux can also be written as $\mathcal{F}_\nu(d) = (\alpha_\star^2/4)\,\mathcal{F}_\nu(R_\star)$. For unresolved objects (such as is the case for almost all stars), we can measure only the flux. If the angular diameter is known, then the measured flux can be converted to the flux at the surface of the object. Reversely, if we know the value for the surface flux from (other) observed quantities we may determine the angular diameter of the object using the measured flux.

2.2 Planck function

We recapitulate properties of the Planck function. The Planck function per frequency unit is given by

$$B_\nu(T) = \frac{2h\nu^3}{c^2} \frac{1}{e^{h\nu/kT} - 1} \tag{2.23}$$

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

$$|B_\nu(T)\, d\nu| = |B_\lambda(T)\, d\lambda| \tag{2.24}$$

we may rewrite the Planck function in wavelength units. We find

$$B_\lambda(T) = \frac{2hc^2}{\lambda^5} \frac{1}{e^{hc/\lambda kT} - 1} \tag{2.25}$$

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

Wien's displacement law

Note that if we write $B_\nu(T) = T^3 q^3 \mathcal{B}(q)$, where $q = \nu/T$, that the function $\mathcal{B}(q)$ is a universal function. This proofs that Planck curves for different temperature will not intersect.

The position of the maximum of the Planck curves follows from *Wien's displacement law*, and can be derived by taking the partial derivative $\partial B_\nu / \partial \nu = 0$. We find

$$h\nu_{\max} = 2.82144\, kT \qquad \text{or} \qquad \frac{\nu_{\max}}{T} = 5.87870 \times 10^{10} \quad \text{hz K}^{-1} \tag{2.26}$$

The peak of B_λ is at

$$\lambda_{\max} T = 0.28979 \quad \text{cm K} \tag{2.27}$$

Note: the maxima of B_ν and B_λ are *not* at the same position in the spectrum.

Rayleigh-Jeans approximation

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

$$B_\nu(T) \simeq \frac{2\nu^2 kT}{c^2} \tag{2.28}$$

The Rayleigh-Jeans approximation is particularly important in radio astronomy, where the condition $h\nu \ll kT$ is most easily met. At these long wavelengths a fixed ration between B_ν and T exists, so that the intensity can be expressed in Kelvin.

Wien's approximation

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

$$B_\nu(T) \simeq \frac{2h\nu^3}{c^2} e^{-h\nu/kT} \tag{2.29}$$

Stefan-Boltzmann's law

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

$$B \equiv \int_0^\infty B_\nu \, d\nu = \frac{\sigma}{\pi} T^4 \tag{2.30}$$

where

$$\sigma = \frac{2\pi^5 k^4}{15 h^3 c^2} = 5.66961 \times 10^{-5} \quad \text{erg cm}^{-2}\,\text{K}^{-4}\,\text{s}^{-1} \tag{2.31}$$

is Stefan-Boltzmann's constant.

2.3 The equation of transfer

In this section we formulate the equation of transfer of radiation through a medium and introduce the macroscopic quantities that play a role in this equation. The transfer equation has

a formal solution which reflects that the specific intensity at each point in the medium can be determined if the source function S_ν and the optical depth τ_ν are known throughout the medium, or – equivalently – the extinction coefficient χ_ν and the emission coefficient η_ν. We mention analytical solutions to simple transfer problems.

Extinction coefficient

If only extinction processes occur along the path-length ds, either as a result of absorption or because of scattering out of the direction of the ray, then the decrease in the specific intensity is proportional to the incident specific intensity, to the path-length ds, and to the properties and number of absorbing/scattering particles. The constant of proportionality is called the extinction coefficient and can be defined in three ways (see eq. 2.34 for an overview).

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

$$dI_\nu(s) = -\chi_\nu(s)\, I_\nu(s)\, ds \tag{2.32}$$

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

$$\chi_\nu(s) = \kappa_\nu(s) + \sigma_\nu(s) \tag{2.33}$$

Here it is implicitly assumed that these processes are independent of each other and that they are additive.

Alternative definitions of the extinction coefficient are

$$\chi_\nu = \chi\prime_\nu\, \rho = \alpha_\nu\, n \tag{2.34}$$

Here $\chi\prime_\nu$ is the *mass extinction coefficient* in cm^2 gr^{-1} and ρ the density of the material medium in gr cm^{-3}. α_ν is the extinction coefficient per particle or *cross section* in cm^2 and n the number density of particles in cm^{-3} that cause the extinction.

Emission coefficient

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

$$dI_\nu(s) = \eta_\nu(s)\, ds \tag{2.35}$$

The constant of proportionality is called the *volume emission coefficient* and has dimensions erg cm^{-3} sec^{-1} hz^{-1} sr^{-1}. η_ν always depends on the properties of the medium, and, in case of scattering from other directions into the ray of light, also on the radiation field. For

completeness we mention that in studies of gaseous nebulae it is custom to use the letter j (so j_ν) to denote the emission coefficient (for instance in Osterbrock & Ferland).

An alternative definition of the emission coefficient is

$$\eta_\nu = \eta\prime_\nu \rho \tag{2.36}$$

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

Extinction and emission: the equation of transfer

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

$$dI_\nu(s) = [\,\eta_\nu(s) - \chi_\nu(s)\, I_\nu(s)\,]\, ds \tag{2.37}$$

or

$$\frac{\mathrm{d}I_\nu}{\mathrm{d}s} = \eta_\nu - \chi_\nu\, I_\nu \tag{2.38}$$

This is the *equation of transfer*.

Optical depth

Let us introduce an elementary *optical depth interval* $d\tau_\nu$ along a path-length ds, such that

$$d\tau_\nu \equiv \chi_\nu(\mathbf{r}, \mathbf{n})\, ds \tag{2.39}$$

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

$$\tau_\nu(D) = \int_0^D \chi_\nu(s)\, ds \tag{2.40}$$

The optical depth is a dimensionless number. The physical meaning of τ_ν is simple. In the absence of emission in the medium the equation of transfer along the beam is given by (see eq. 2.32 and/or 2.38)

$$\frac{\mathrm{d}I_\nu}{\mathrm{d}\tau_\nu} = -I_\nu \tag{2.41}$$

This yields

$$I_\nu(D) = I_\nu(0)\, e^{-\tau_\nu(D)} \tag{2.42}$$

and shows that $\tau_\nu(D)$ is the exponential decline parameter that determines what remains of a beam that has passed through a layer of thickness D in which extinction processes occur. How far can photons penetrate in this layer? The chance that an incident photon travels an optical depth $\tau_\nu(s)$ in the layer (for $s < D$) before it is absorbed or scattered is $p(\tau_\nu(s)) =$

$\exp(-\tau_\nu(s))$. As the average of a quantity x that has a probability distribution $p(x)$ is given by

$$< x >= \frac{\int_0^\infty x\, p(x)\, dx}{\int_0^\infty p(x)\, dx} \tag{2.43}$$

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

$$< \tau_\nu >= \frac{\int_0^\infty \tau_\nu\, e^{-\tau_\nu}\, d\tau_\nu}{\int_0^\infty e^{-\tau_\nu}\, d\tau_\nu} = 1 \tag{2.44}$$

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

Mean free path of the photon

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

$$\ell_\nu = \frac{< \tau_\nu >}{\chi_\nu} = \frac{1}{\chi_\nu} = \frac{1}{\chi'_\nu \rho} = \frac{1}{\alpha_\nu\, n}, \tag{2.45}$$

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

Source function

The *source function* is defined as

$$S_\nu \equiv \frac{\eta_\nu}{\chi_\nu} \tag{2.46}$$

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

From the definition of optical depth, eq. (2.39), it follows that $\eta_\nu\, ds = (\eta_\nu/\chi_\nu)\, \chi_\nu\, ds = S_\nu\, d\tau_\nu$. So, the source function is essentially the amount of emission per unit optical depth.

Formal solution of the transfer equation

We define the optical depth scale such that for distance z increasing in the direction to the observer, τ_ν is decreasing, i.e.

$$d\tau_\nu(z) = -\chi_\nu(z) dz \tag{2.47}$$

This implies that for an observer at $z = \infty$, the optical depth $\tau_\nu(\infty) = 0$. Adopting this optical depth scale, and using the source function, the transfer equation (Eq. 2.38) may be written in its standard form

$$\frac{\mathrm{d}I_\nu(\tau_\nu)}{\mathrm{d}\tau_\nu} = I_\nu(\tau_\nu) - S_\nu(\tau_\nu) \tag{2.48}$$

For known source function S_ν, eq. (2.48) has a formal solution. To find this solution we bring I_ν to the left hand side, and multiply both sides by the integrating factor $\exp(-\tau_\nu)$. This yields

$$\left[\frac{\mathrm{d}I_\nu}{\mathrm{d}\tau_\nu} - I_\nu\right] e^{-\tau_\nu} = \frac{\mathrm{d}}{\mathrm{d}\tau_\nu}(I_\nu\, e^{-\tau_\nu}) = -S_\nu\, e^{-\tau_\nu} \tag{2.49}$$

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

$$I_\nu(\tau_1) = I_\nu(\tau_2)\, e^{-(\tau_2-\tau_1)} + \int_{\tau_1}^{\tau_2} S_\nu(t_\nu)\, e^{-(t_\nu-\tau_1)}\, dt_\nu \tag{2.50}$$

For $\tau_1 = 0$ and $\tau_2 = \tau_\nu$, this reduces to

$$I_\nu(0) = I_\nu(\tau_\nu)\, e^{-(\tau_\nu)} + \int_0^{\tau_\nu} S_\nu(t_\nu)\, e^{-t_\nu}\, dt_\nu \tag{2.51}$$

Solution of the transfer equation for a homogenous slab

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

$$I_\nu(0) = I_\nu(\tau_\nu)\, e^{-\tau_\nu} + S_\nu\left[1 - e^{-\tau_\nu}\right] \tag{2.52}$$

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

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

$$I_\nu(0) \simeq I_\nu(\tau_\nu) + (S_\nu - I_\nu(\tau_\nu))\, \tau_\nu \tag{2.53}$$

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

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

$$I_\nu(0) \simeq S_\nu \tag{2.54}$$

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

We rewrite the solution of eq. (2.52) to

$$I_\nu(0) = I_\nu(\tau_\nu) + (S_\nu - I_\nu(\tau_\nu)) \left[1 - e^{-\tau_\nu}\right] \tag{2.55}$$

Note that $I_\nu(0) < I_\nu(\tau_\nu)$ if $S_\nu < I_\nu(\tau_\nu)$, and reversely that $I_\nu(0) > I_\nu(\tau_\nu)$ if $S_\nu > I_\nu(\tau_\nu)$. This is an important result. It shows the principle of the formation of absorption *casu quo* emission lines in a plane-parallel medium. Let us identify the incident intensity $I_\nu(\tau_\nu)$ with the intensity that is emitted by the background source. The homogeneous layer corresponds to the interstellar medium, i.e. the region in which the spectral lines are formed. Is the line source function $S_\nu < I_\nu(\tau_\nu)$ then an absorption line will form, i.e. $I_\nu(0) < I_\nu(\tau_\nu)$. Is the line source function $S_\nu > I_\nu(\tau_\nu)$ then an emission line will form, i.e. $I_\nu(0) > I_\nu(\tau_\nu)$. Is the line source function equal to the intensity emitted by the continuum, i.e. $S_\nu = I_\nu(\tau_\nu)$, then no line will form, i.e. $I_\nu(0) = I_\nu(\tau_\nu)$.

2.4 Thermodynamic Equilibrium

The most important parameters describing the material medium are the mass density $\rho(\mathbf{r})$ and the temperature $T(\mathbf{r})$. The essential problem in the study of astrophysical media is to understand the interaction between the material medium and the radiation field. The description of the coupling between the gas or dust and the radiation field can be simplified considerably if the medium and the radiation are in some type of equilibrium.

The most strict form of equilibrium is *thermodynamic equilibrium* (TE). In TE the medium is homogeneous and at rest. There are no gradients. Each process is in a microscopic equilibrium with its reverse process, i.e. all processes are in *detailed balance*. The radiation field is given by the Planck function, which only depends on T, so

$$I_\nu = I_\nu(T) \equiv B_\nu(T), \tag{2.56}$$

and both the thermal velocity distribution of the particles – given by the *Maxwell velocity distribution* – and the distribution of particles over excitation and ionization states – given by the *Boltzmann and Saha distributions* – depend only on ρ and T.

However, the simple fact that we receive photons of the objects in which we are interested tells us that these media can not be in a state of TE. Because photons escape from these media it must be so that significant gradients are present in the quantities describing the medium.

If we can not assume that the medium may be characterized as a whole with one value for ρ and one value for T, but if we are allowed to describe the state of the material medium *locally* using only the value of $\rho(\mathbf{r})$ and $T(\mathbf{r})$, then we refer to the situation as being in *local thermodynamic equilibrium* (LTE). In LTE all atomic processes are still in detailed balance. However, the radiation field is *not* in equilibrium, but follows from the solution of the equation of transfer.

In those cases that even LTE is not valid, the medium is (per definition) in a state of *non local thermodynamic equilibrium* (NLTE). This is so when (at least) one microscopic process is not in detailed balance. The population of (at least) two excitation/ionization levels will deviate from their LTE values. The Boltzmann and Saha distributions no longer hold, and one requires a *statistical equilibrium* analysis to solve for the distributions of excitation and ionization states.

In this section we mainly focus on a description of TE.

Kirchhoff's law & the Kirchhoff-Planck relation

In a medium that is in TE there are no gradients, i.e. the specific intensity is homogeneous, isotropic and time independent. In that case eq. (2.38) implies that for all rays, for all frequencies, at all times

$$\eta_\nu = \chi_\nu \, I_\nu \tag{2.57}$$

This is *Kirchhoff's law*. Combining this law with eq. (2.56) results in

$$\eta_\nu = \chi_\nu \, B_\nu(T) \tag{2.58}$$

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

In summary: in TE $I_\nu = B_\nu = S_\nu$. As the radiation field is isotropic, also $I_\nu = J_\nu$. Because there are no gradients, the flux $\mathcal{F}_\nu = 0$.

Maxwellian velocity distribution and kinetic gas-temperature

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

$$f(\mathbf{v}) \, dv_x dv_y dv_z = \left(\frac{m}{2\pi kT}\right)^{3/2} \exp\left[-m(v_x^2 + v_y^2 + v_z^2)/2kT\right] \; dv_x dv_y dv_z \tag{2.59}$$

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

$$f(v_x) \, dv_x = \frac{1}{\sqrt{\pi}} \left(\frac{m}{2kT}\right)^{1/2} \exp\left[-mv_x^2/2kT\right] \; dv_x \tag{2.60}$$

The most probable velocity in the x direction is $v_x = 0$. The root-mean-square velocity in this direction is $\langle v_x^2 \rangle^{1/2} = (kT/m)^{1/2}$.

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

$$f(v) \, dv = \left(\frac{m}{2\pi kT}\right)^{3/2} \exp\left[-mv^2/2kT\right] \; 4\pi v^2 \, dv \tag{2.61}$$

The dimensions of $f(v)$ are cm^{-1} sec and the integral of $f(v)$ is unity, similar to $f(\mathbf{v})$ and $f(v_x)$. It is not a Gauss distribution, but shows a "tail" as a result of the v^2 term. The most probable speed in this distribution, i.e. the one for which $f(v)$ reaches a maximum, is $v = (2kT/m)^{1/2}$. The mean speed is

$$\langle v \rangle = \left(\frac{8kT}{\pi m}\right)^{1/2} = 14.551 \left(\frac{T}{10^4 A}\right)^{1/2} \quad \mathrm{km\,s^{-1}}, \tag{2.62}$$

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

$$\langle v^2 \rangle^{1/2} = \left(\frac{3kT}{m}\right)^{1/2} = 15.793 \left(\frac{T}{10^4 A}\right)^{1/2} \quad \mathrm{km\,s^{-1}}. \tag{2.63}$$

This implies an average thermal energy of the particle

$$\frac{1}{2} m \langle v^2 \rangle = \frac{3}{2} kT. \tag{2.64}$$

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

The average relative speed of two particles of mass m_1 and m_2 of a distribution of gas particles that is given by the Maxwell velocity distribution can be computed by substituting the reduced mass $\mu = m_1\, m_2/(m_1 + m_2)$ for the mass m in Eq. (2.61) and yields a result similar to Eq. (2.62), namely

$$\langle v \rangle = \left(\frac{8kT}{\pi \mu}\right)^{1/2}, \tag{2.65}$$

From this result one may derive an expression for the average relative speed of hydrogen atoms with respect to particles with a mass $A\, m_{\rm amu}$. We find

$$\langle v \rangle = 0.145\, T^{1/2} (1 + 1/A)^{1/2} \quad \mathrm{km\,s^{-1}}. \tag{2.66}$$

So, the average relative speed of two hydrogen atoms (A =1) in interstellar gas at a temperature of $T = 70$ K is $\langle v \rangle$ = 1.7 km s^{-1}.

If the Maxwellian velocity distribution is valid one speaks of *kinetic equilibrium* (KE). For this condition to be valid a particle that is ejected into the gas (for instance, an electron that is ejected from an ion after a photo-ionization or a collisional ionization) needs to experience a large number of elastic collisions (for the electron that we use as an example these are usually collisions with other free electrons) before it suffers a non-elastic collision. KE implies that the medium has a unique *kinetic temperature*.

Level populations and Boltzmann distribution

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

$$\frac{n_j}{n_i} = \frac{g_j}{g_i} \exp\left[-(E_j - E_i)/kT\right] = \frac{g_j}{g_i} \exp\left[-h\nu_{ij}/kT\right]. \tag{2.67}$$

Here n_i is the number density per cm^3 in level i; g_i is the *statistical weight*; E_i the *excitation energy* in erg, and $\nu_{ij} = \Delta E_{ij}/h$ the frequency in Hz corresponding to the energy difference $\Delta E_{ij} = E_j - E_i$. Limiting cases are

$$h\nu_{ij} \gg kT \rightarrow n_j = 0 \tag{2.68}$$

$$h\nu_{ij} \ll kT \rightarrow n_j/n_i = g_j/g_i \tag{2.69}$$

It is customary that within each ionization stage the excitation energy is measured from the ground state up (see figure 2.5). The ionization energy is also measured per ionization stage from the ground state up. Usually one does not give the energy difference between levels in erg, but in electron volts (eV) or kayser (cm^{-1}). If we adopt the first unit one speaks of *excitation potential* . In this case E_j[erg] = 1.602192 10^{-12} E_j[eV] and ΔE_{ij}[eV] = 12 398.54 / λ_{ij}[Å]. If we adopt the last unit the wavelength of the transition can trivially be recovered using λ_{ij}[Å] = $10^8/(E_j$[cm^{-1}] - E_i[cm^{-1}]). Figure 5.1 gives the excitation energy in Kelvin, i.e. it provides E_j/k.

The Boltzmann equation can also be written to yield the excitation fraction

$$\frac{n_i}{N} = \frac{g_i}{U(T)} \exp\left[-E_i/kT\right], \tag{2.70}$$

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

$$U(T) \equiv \sum_i g_i \exp\left[-E_i/kT\right]. \tag{2.71}$$

Ionization ratios and Saha distribution

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

$$\begin{aligned} \frac{N_I}{N_{I+1}} &= n_\mathrm{e} \frac{U_I}{2U_{I+1}} \left(\frac{h^2}{2\pi m_\mathrm{e} kT}\right)^{3/2} \exp\left[E_I/kT\right] \\ &\equiv n_\mathrm{e} \tilde{\Phi}_I(T) \end{aligned} \tag{2.72}$$

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

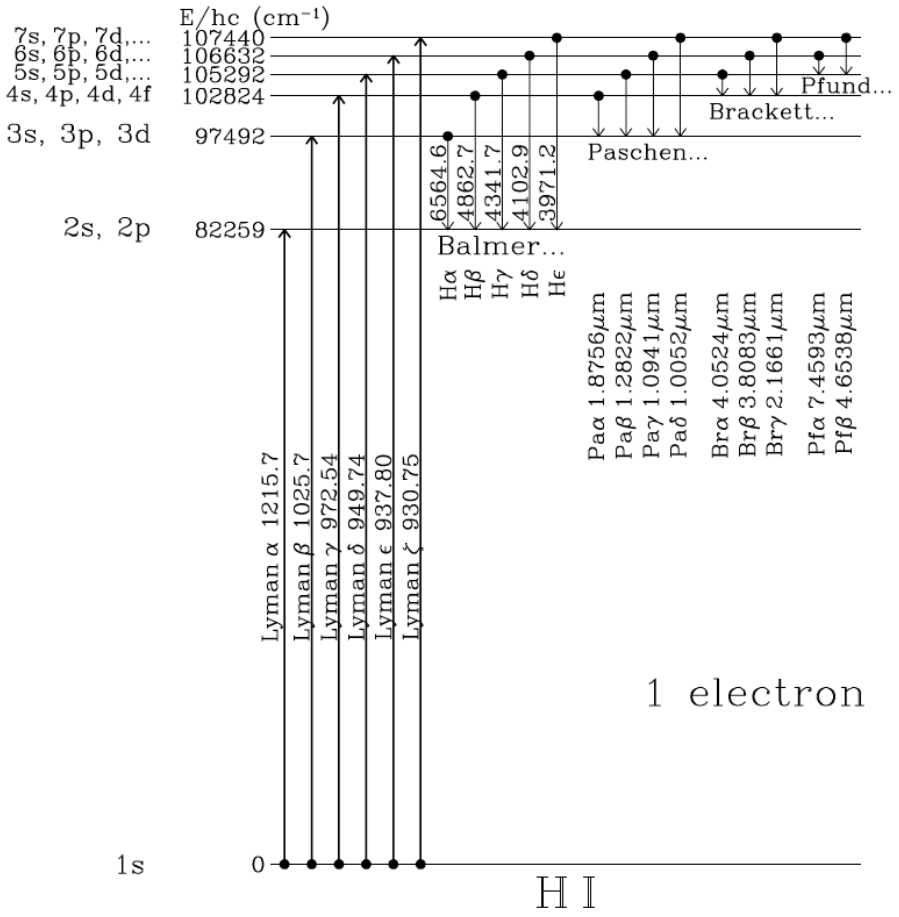


Figure 2.5: *Schematic picture of the energy level diagram or Grotrian diagram of hydrogen. The excitation energy of the levels is measured from the ground state up and is given in units of* cm^{-1} *or kayser. From: Bruce Draine,* Physics of the interstellar and intergalactic medium.

If the temperature decreases at constant density (such that n_e remains more or less the same) the exponential term in the Saha ionization equation will make that N_I/N_{I+1} increases, i.e. the medium *recombines*. Essentially, the root-mean-square speed (Eq. 2.63) of the Maxwellian velocity distribution decreases, lowering the number of free electrons that can cause ionization. If the density decreases at constant temperature, n_e (more or less proportional to ρ) will decrease and therefore also N_I/N_{I+1}, i.e. the medium *ionizes*. Essentially, the lower electron density lowers the probability that a free electron is captured by an ion.

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

$$\frac{n_{0,I}}{n_{0,I+1}} = n_\mathrm{e} \frac{g_{0,I}}{2g_{0,I+1}} \left(\frac{h^2}{2\pi m_\mathrm{e} kT} \right)^{3/2} \exp\left[E_I/kT\right] \tag{2.73}$$

Finally, we may express the particle density n_i in terms of the particle density of the ground level of ion $I+1$ and the electron density by substitution of eq. (2.67) in (2.73). This gives

$$\begin{aligned} n_{i,I} &= n_{0,I+1} n_{\rm e} \frac{g_{i,I}}{2g_{0,I+1}} \left(\frac{h^2}{2\pi m_{\rm e} kT} \right)^{3/2} \exp\left[(E_I - E_i)/kT \right] \\ &\equiv n_{0,I+1} n_{\rm e}\, \Phi_{iI}(T) \end{aligned} \tag{2.74}$$

This is the *equation of state* (EOS) of a gas in TE or LTE. In the latter case one should use the local values for temperature and density. The equation can also be used to *define* LTE particle densities in a NLTE medium. Deviations from LTE can thus be described by the *NLTE departure coefficient*

$$b_{i,I} \equiv \frac{n_{i,I}}{n_{i,I}^{\rm LTE}} \tag{2.75}$$

in which $n_{i,I}^{\rm LTE}$ is given by eq. (2.74). Henceforth we will denote LTE values of the level populations with a superscript $*$ (and not with LTE). However, we will do so only if this is required in the context of the discussion (if not, no superscript is used).

2.5 Local Thermodynamic Equilibrium

Local Thermodynamic Equilibrium or LTE implies that the conditions of TE apply at a specific location in the medium, where the local temperature and density are $T(\mathbf{r})$ and $\rho(\mathbf{r})$. This does not include the TE description of the radiation field. In LTE one must solve the equation of transfer Eq. 2.48. LTE is a good description of the medium for one or both of the following conditions:

- Densities are so high that collisions dominate the electron transitions between levels. These collisions couple the level populations to the local medium (as the rate of collisions depends on $T(\mathbf{r})$ and $\rho(\mathbf{r})$ only).

- The optical depth in all transitions of an atom, ion, or molecule are so high that photons are locally truly absorbed or truly emitted. In that case the level populations depend on the radiation field that is locally generated, i.e. on the local Planck function $B_\nu(T(\mathbf{r}))$.

In LTE the temperature and density may vary gradually with location in the medium, however, densities are always high enough to ensure that the state of the gas is determined locally. If, for instance, photons created at some remote location reach $\mathbf{r}$ through a series of scatterings, these should not contribute in a significant way to the local radiation field. If this would be the case, these photons would likely 'report' on conditions elsewhere, where the temperature and density may be very different from those at $\mathbf{r}$, and introduce departures from LTE.

2.6 Temperature definitions

Based on the above discussions we may define three characteristic temperatures that are often used in the study of the ISM. The first two are related to the Planck function; the next two to the equations of Boltzmann and Saha.

Brightness temperature

The *brightness temperature* T_b is the temperature for which the Planck function, at the frequency at which is measured, reproduces the observed specific intensity

$$I_\nu^\mathrm{obs} = B_\nu(T_\mathrm{b}) \tag{2.76}$$

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

$$T_\mathrm{b} \equiv \frac{h\nu/k}{\ln\left[1 + 2h\nu^3/c^2 I_\nu^\mathrm{obs}\right]} \simeq \frac{c^2}{2\nu^2 k} I_\nu^\mathrm{obs} \tag{2.77}$$

In general T_b will depend on frequency. Only if the source emits a blackbody spectrum T_b will be the same for all frequencies.

If we consider a medium at a constant temperature T that emits according to the Planck function, the solution of the equation of transfer (eq. 2.52) in the Rayleigh-Jeans limit is

$$T_\mathrm{b}(0) = T_\mathrm{b}(\tau_\nu)\, e^{-\tau_\nu} + T\left[1 - e^{-\tau_\nu}\right] \qquad \forall\ \frac{h\nu}{kT} \ll 1 \tag{2.78}$$

where $T_\mathrm{b}(\tau_\nu)$ is the brightness temperature of the radiation field that is incident at the far side of the medium. For large optical depth in the medium the brightness temperature will be equal to the temperature of the material, i.e. $T_\mathrm{b}(0) = T$.

Excitation temperature

The excitation temperature is determined by comparison of observed level populations of an atom, ion, or molecule, using the Boltzmann equation (eq 2.67). Obviously, the measured populations need not be in LTE. To conserve a description of the measured population ratios in terms of the Boltzmann equation, we introduce the excitation temperature T_ext. This temperature is given by

$$\left(\frac{n_j}{n_i}\right)^\mathrm{obs} = \frac{b_j}{b_i}\frac{n_j^*}{n_i^*} = \frac{b_j}{b_i}\frac{g_j}{g_i}\exp\left[-h\nu_{ij}/kT\right] \equiv \frac{g_j}{g_i}\exp\left[-h\nu_{ij}/k\,T_\mathrm{ext}\right] \tag{2.79}$$

If $b_u/b_l = 1$, such as is the case in TE and LTE, then $T = T_\mathrm{ext}$. An interesting case is when $T_\mathrm{ext} < 0$, so that $n_j/n_i > g_j/g_i$. This means that an inversion of the level populations with respect to TE or LTE occurs, which is a condition for maser amplification.

Ionization temperature

In analogy to the excitation temperature, the *ionization temperature* is the temperature for which an observed ionization ratio fulfills the Saha equation (eq. 2.72)

$$\left(\frac{N_I}{N_{I+1}}\right)^{\mathrm{obs}} \equiv n_\mathrm{e}\frac{U_I}{2U_{I+1}}\left(\frac{h^2}{2\pi m_\mathrm{e} k\,T_\mathrm{ion}}\right)^{3/2}\exp\left[E_I/k\,T_\mathrm{ion}\right] = n_\mathrm{e}\tilde{\Phi}_I(\,T_\mathrm{ion}) \qquad (2.80)$$

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

Exercise 2.1

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

a) Compute the radius of the moon.

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

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

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

Malcolm Smith, Hertfordshire, England

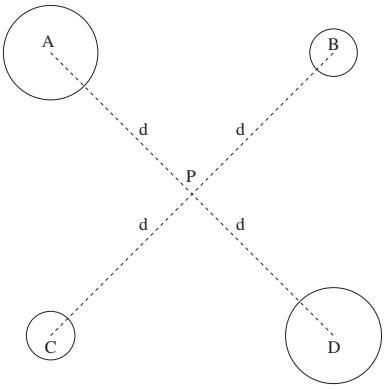
Published: Sunday, October 1, 2000

October 2000

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

Exercise 2.2

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



a) Give the mean intensity at P.

b) Give the monochromatic energy density at P.

c) Give the flux at P.

Exercise 2.3

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

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

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

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

Exercise 2.4

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

Exercise 2.5

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

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

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

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

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

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

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

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

Exercise 2.6

a) Give the derivation of eq. (2.25) using eq. (2.23), and the relation (2.24)

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

c) Derive eq. (2.30) and the constant (2.31), given that

$$\int_0^\infty \frac{x^3}{e^x - 1}\, dx = \frac{\pi^4}{15} \tag{2.81}$$

Exercise 2.7

a) The number of photons per cm^3 per Hz in a Planck radiation field is $n(\nu) = u_\nu / h\nu$, where u_ν (see eq. 2.11) is the energy density of the Planck radiation field. Show that the total number of photons per cm^3 in the Planck radiation field

$$n_{\rm tot} = \frac{16\pi\zeta(3)k^3}{c^3h^3}\, T^3 \simeq 20\, T^3 \tag{2.82}$$

given that

$$\int_0^\infty \frac{x^2}{e^x - 1}\, dx = 2\,\zeta(3) \tag{2.83}$$

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

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

Exercise 2.8

a) Derive eq. (2.61) from (2.59).

b) Show that both distributions are normalized. Remember that

$$\int_{-\infty}^{+\infty} e^{-x^2/a^2}\, dx = a\sqrt{\pi} \qquad \text{and} \qquad \int_0^{+\infty} x^2 e^{-x^2/a^2}\, dx = a^3\sqrt{\pi}/4 \tag{2.84}$$

c) Show that the mean particle energy is given by

$$< \frac{1}{2}mv^2 >= \frac{3}{2}kT \tag{2.85}$$

if we give that

$$\int_0^{+\infty} x^4 e^{-x^2/a^2}\, dx = \frac{3}{8}a^5\sqrt{\pi} \tag{2.86}$$

d) Show that the most probable velocity of eq. (2.60) is given by $v_x = 0$ and of eq. (2.61) by $v = (2kT/m)^{1/2}$.

Exercise 2.9

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

$$n_{\rm e}(N_{\rm H}, T) = \left[\left(4N_{\rm H}\tilde{\Phi}_{\rm H}(T) + 1\right)^{1/2} - 1\right] \Big/ 2\tilde{\Phi}_{\rm H}(T), \tag{2.87}$$

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

Exercise 2.10

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

The temperature at the time of recombination can be estimated through the use of the Saha equation (2.72) for neutral and ionized hydrogen, where we use $U_1 \equiv U_{\rm I} = 2$ and $U_2 \equiv U_{\rm II} = 1$. The ionisation energy of hydrogen is $E_{\rm I}$ = 13.6 eV. We define q to be the fraction of hydrogen atoms that are ionised, i.e.

$$q = \frac{N_{\rm II}}{N_{\rm I} + N_{\rm II}}, \tag{2.88}$$

hence

$$\frac{N_{\rm II}}{N_{\rm I}} = \frac{q}{1-q}, \tag{2.89}$$

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

$$z = \frac{1}{R} - 1. \tag{2.90}$$

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

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

c) Use Saha equation (2.72) to derive an expression that links the scale factor R to the ionization fraction q, for the known present-day quantities $n_{\mathrm{b},\circ}$ and $T_\circ$.

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

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

Exercise 2.11

a) Does the brightness temperature of a radio source depend on its distance?

b) Can one measure the brightness temperature of a point source (i.e. an object that is not spatially resolved) such as a star? Can one measure T_{b} for an extended source, such as a nebula, if this source is not in TE?

3

Atomic structure

Atomic structure describes the organization of electrons in various shells and subshells. This chapter briefly summarizes the energy-level structure of atoms and ions, together with the nomenclature for referring to those levels, and may be regarded as reference material. Many textbooks discuss this notation; here we particularly use parts of the clear and concise summary of Draine (2011), who often relies on the excellent book *Physics of Atoms and Molecules* by Bransden & Joachain (2003).

3.1 Single electron orbitals

Hydrogen is the most common element in the universe and it, obviously, plays an important role in the study of interstellar gas. It is also the simplest atom with only a single electron orbiting the nucleus. It is therefore with hydrogen that a study of atomic structure must begin.

For a one-electron system in a Coulomb field the orbital of the electron is characterized by an integer principal quantum number $n = 1, 2, 3, 4, ...$ and an orbital angular momentum quantum number ℓ, that can assume integer values $0 \leq \ell < n$. Electrons with the same value of n are referred to as being in the same *shell*. Electrons with the same value of n and ℓ are referred to as being in the same *subshell*. Though ℓ can thus take values $0, 1, 2, 3, ...$ it is custom to use lowercase letters to denote these values, as these were first used by spectroscopists before quantum mechanics was known. In the proper order these lowercase letters are $\ell = s, p, d, f, g, ...$ The first four letters refer to 'sharp', 'principal', 'diffuse', and 'fundamental'. The actual magnitude of the orbital angular momentum $\mathbf{L}$ is denoted L and is related to the quantum number ℓ by

$$|\mathbf{L}| = L = \sqrt{\ell\,(\ell+1)}\,\hbar. \tag{3.1}$$

In single electron orbitals the value of ℓ has no effect on the energy of the electron state (but see *fine structure* below); only n does to any appreciable extent. The energies the electron

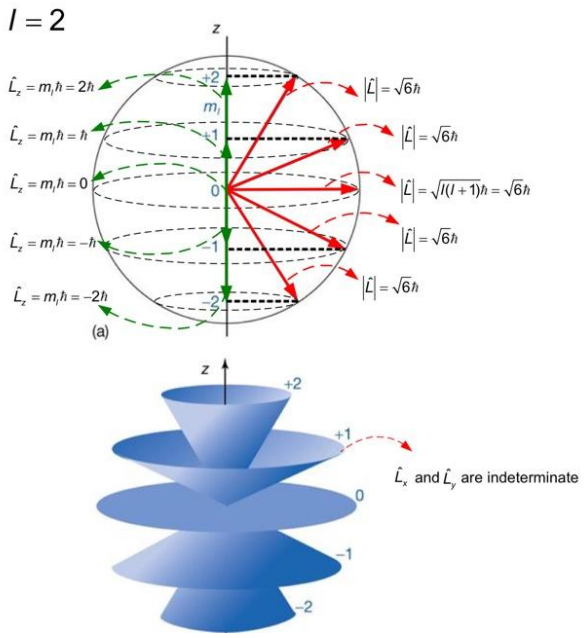


Figure 3.1: *Quantization of the orientation of the orbital angular momentum* $\mathbf{L}$ *in space.* $\mathbf{L}$ *takes on only those orientations for which the* z *component* L_z *has values given by integer multiples of* $\hbar$*. For a unique value of* L_z *the azimuthal angle is completely undetermined, i.e.* L_x *and* L_y *are indeterminate. Like all classical representations of quantum mechanical results, the visualization can be misleading and must be used with caution. For example, it may lead to experimentally meaningless questions such as, 'In what direction is the angular momentum vector* really *pointing at such-and-such instant?' From: chegg.com.*

may assume are

$$E_n = -\frac{\mu\, c^2\, Z^2\, \alpha^2}{2}\frac{1}{n^2} = -\frac{13.595\,\mathrm{eV}}{n^2}, \tag{3.2}$$

where μ is the reduced mass of the nucleus-electron system, that is,

$$\mu = \frac{m_\mathrm{N}\, m_\mathrm{e}}{(m_\mathrm{N} + m_\mathrm{e})}, \tag{3.3}$$

where m_N is the mass of the nucleus, m_e the mass of the electron, and $\alpha = 7.297 \times 10^{-3} \simeq 1/137$ is the dimensionless fine structure constant. In the last equality of Eq. 3.2, we consider hydrogen, which has a nuclear charge $Z = 1$.

In addition to the quantum numbers n and ℓ, there is a third quantum number characterizing the orbital: m_ℓ, the projection of the orbital angular momentum onto an axis that is fixed in one direction, L_z; see Fig. 3.1. The choice of z as the relevant direction is arbitrary. Thus m_ℓ can take on $2\ell + 1$ different values: $-\ell$, ... , -1, 0, $+1$, ..., $+\ell$. The actual magnitude of L_z is related to m_ℓ by

$$L_z = m_\ell\, \hbar. \tag{3.4}$$

This limitation on the direction of $\mathbf{L}$ is often called *space quantization*.

The name m_ℓ derives not from theory (which relates it to L_z), but from experiment. It was found that when a gas-discharge tube was placed in a magnetic field, the spectral lines were split into several very closely spaced lines. This splitting, discovered by the Dutch physicist Pieter Zeeman (1865–1943) and known as the Zeeman effect, implies that the energy levels must be split, and thus that the energy of a state depends not only on n but also on m_ℓ when a magnetic field is applied – hence the name magnetic quantum number or magnetic angular quantum number.

Finally, the electron possesses a spin quantum number s that is the same for all electrons and equal to $s = 1/2$. This causes it to have a spin angular momentum $\mathbf{S}$ that has a magnitude

$$|\mathbf{S}| = S = \sqrt{s\,(s+1)}\,\hbar = \frac{\sqrt{3}}{2}\,\hbar. \tag{3.5}$$

The z component of the spin angular momentum is

$$S_z = m_s\,\hbar, \tag{3.6}$$

where the two possible spin quantum numbers m_s can take the values $+1/2$ (called 'spin up') and $-1/2$ (called 'spin down'). Thus a given pair of quantum numbers $n\ell$ actually refers to $2(2\ell+1)$ distinct electronic wave functions. We include m_s, but not s, in our list of quantum numbers as s is the same for all electrons. A given n refers to

$$g = \sum_{0}^{n-1} 2(2\ell+1) = 2n^2 \tag{3.7}$$

distinct wave functions or quantum states, where g is the degeneracy or statistical weight of a hydrogen level.

The existence of m_s did not come out of Schrödinger's original theory (as did n, ℓ, and m_ℓ). It was Dirac (1902–1984) who introduced it in the theoretical description of the atom. The first hint that m_s was needed, however, came from experiment. A careful study of the spectral lines of hydrogen showed that each actually consisted of two (or more) very closely spaced lines even in the absence of an external magnetic field. It was first hypothesized that this tiny splitting of energy levels, called *fine structure*, was due to a magnetic field produced by the atom itself. We can see how it occurs by putting ourselves in the reference frame of the electron, in which case we see the nucleus revolving around us as a moving charge or electric current that produces a magnetic field. The electron has an intrinsic magnetic dipole moment of which the z-component is proportional to m_s, and hence the energy of the single electron state will split into two closely spaced energy levels (for $m_s = +1/2$ and $m_s = -1/2$)[1]. In

[1]The z-component of the electron magnetic dipole moment is $(\mu_s)_z = -g_s\mu_B m_s$, where μ_B is the Bohr magneton physical constant and $g_s \sim 2$ is the spin g-factor. Note the negative constant; i.e., the magnetic moment is antiparallel to the spin angular momentum.

general, fine structure is therefore said to be due to a *electron spin-orbit interaction*[2].

The magnitude of the energy difference $\Delta E \propto \mathbf{L} \cdot \mathbf{S} \propto \alpha^2 Z^4$ eV, where $\alpha \simeq 1/137$ is the fine structure constant and Z is the nuclear charge, i.e., the atomic number. We find that $\Delta E \sim 5 \times 10^{-5}\, Z^4$ eV. Transitions among fine-structure levels occur therefore at near- (heavy atoms, i.e. large nuclear charge Z) to far-IR wavelengths (light atoms).

Today we no longer consider the picture of a spinning electron as legitimate. We cannot even view an electron as a localized object, much less a spinning one. What remains is that the electron can have two different states due to some intrinsic property that behaves like an angular momentum, and we still call this property 'spin'.

3.2 Multiple electron orbitals

Now we discuss more complex atoms, those that contain more than one electron. Their energy levels are not the same as in the H atom, since the electrons interact with each other as well as with the nucleus. Each electron in a complex atom still occupies a particular state characterized by the same quantum numbers n, ℓ, m_ℓ, and m_s. For atoms with more than one electron, the energy levels depend on both n and ℓ. When an atom or ion has more than one electron, the *Pauli exclusion principle* forbids two electrons from sharing the same orbital. Therefore, there can be at most $2(2\ell + 1) = 4\ell + 2$ electrons in a given *subshell* $n\ell$: s subshells can contain at most 2 electrons, p subshells can contain at most 6 electrons, and d subshells can contain up to 10 electrons.

The orbitals, in order of increasing energy, are $1s$, $2s$, $2p$, $3s$, $3p$, $4s$, $3d$, $4p$, $5s$, and so on. Thus atomic carbon, C I, with 6 electrons, has a ground state *configuration* with 2 electrons in the $1s$ subshell, 2 electrons in the $2s$ subshell, and the remaining 2 electrons in the $2p$ subshell. The number of electrons in each subshell is designated by a superscript: the ground state configuration for neutral carbon is written as $1s^2\, 2s^2\, 2p^2$. Neutral sodium, with 11 electrons, has ground state configuration $1s^2\, 2s^2\, 2p^6\, 3s$.

Spectroscopic terms

If an orbital has more than one and less than $4\ell + 1$ electrons (for the np subshell, this means 2, 3, or 4 elections), then there will be more than one way in which the orbital and spin angular momentum vectors of the electrons in the partially filled shell can add. In the so-called LS-coupling or Russell-Saunders coupling approximation, the orbital angular momentum of

[2]But not for hydrogen. Though we don't aim to delve too deeply into atomic physics, for hydrogen fine splitting is dominated by a relativistic effect. This is a special feature of hydrogenic atoms. A discussion of this effect is however beyond the scope of the lecture notes. For many-electron atoms it is the spin-orbit effect which is responsible for the fine structure splitting.

multiple electrons add (vectorially) to give a total orbital angular momentum L, and the individual spin angular momenta similarly add (vectorially) to give a total spin angular momentum S. Each allowed (L, S) combination – the wave functions must of course obey the Pauli exclusion principle – is referred to as a *term*. The total angular momentum L for more than one electron follows the same alphabetic character notation as for the single electron case, but is expressed in upper case letters. For example, $L = 0$ is denoted as S (do not get confused with the total spin S), and the higher values $L = 1, 2, 3, 4$ are P, D, F, G. The LS-coupling designation of an atomic state is conventionally expressed as

$$^{2S+1}\mathcal{L}, \tag{3.8}$$

where $\mathcal{L} = S, P, D, F, G$, etcetera. The number of possible quantum states in a term is $(2S + 1)(2L + 1)$.

Let us again take a closer look at neutral carbon. The outer $n\ell$ subshell, $2p^2$, contains two electrons. Each of the p elections has orbital quantum number $\ell = 1$ and spin quantum number $s = 1/2$. With three possible values of $m_\ell = -1, 0, 1$, and two possible values of $m_s = -1/2, +1/2$, there are $3 \times 2 = 6$ possible one-electron states. The Pauli exclusion principle stipulates that both electrons cannot share the same one-electron state, giving $(6 \times 5)/2 = 15$ possible states for the two indistinguishable electrons. We may readily identify these 15 states[3].

1. Both electrons could have $m_\ell = 1$, giving a total orbital angular momentum in the z-direction of $L_z = 2\hbar$. This would require that one electron be spin up and one spin down, so that $S = 0$, as otherwise the Pauli principle would be violated. Having $L_z = 2\hbar$ requires $L \geq 2$. For two $\ell = 1$ orbitals, the maximum possible value of $L = 2$. Thus it is evident that one of the allowed terms has $S = 0$ and $L = 2$, i.e. ^1D. This accounts for $(2S + 1)(2L + 1) = 1 \times 5 = 5$ of the 15 possible quantum states.

2. Both electrons could have $m_s = 1/2$, such that $S = 1$. One electron could have $m_\ell = 1$ and one have $m_\ell = 0$, so that $L_z = 1\hbar$ is possible, requiring this state to have $L \geq 1$. We have seen above that the only way to have $L > 1$ is to have $S = 0$; therefore, this term must have $L = 1$. With degenerate $(2S + 1)(2L + 1) = 3 \times 3$, this ^3P term accounts for 9 quantum states.

3. To complement the 15 quantum state there is one remaining term, with $S = 0$ and $L = 0$ it is ^1S.

Different terms – for an np^2 configuration, the three possible terms ^1D, ^3P, and ^1S – will differ in energy by a significant fraction of the total binding energy of the electrons in the partially filled subshells. For atoms and low-ionization ions, the energy difference between the different terms of the ground state configuration will be of order a few eV. This is shown in Fig. 3.2 for the six electron systems N II and O III, that have an identical spectroscopic term structure as C I. Note that the level energies or excitation energies are given in Kelvin, for which holds that 1 eV = 11604.5 K.

[3]See for instance `http://en.wikipedia.org/wiki/Term_symbol` for details.

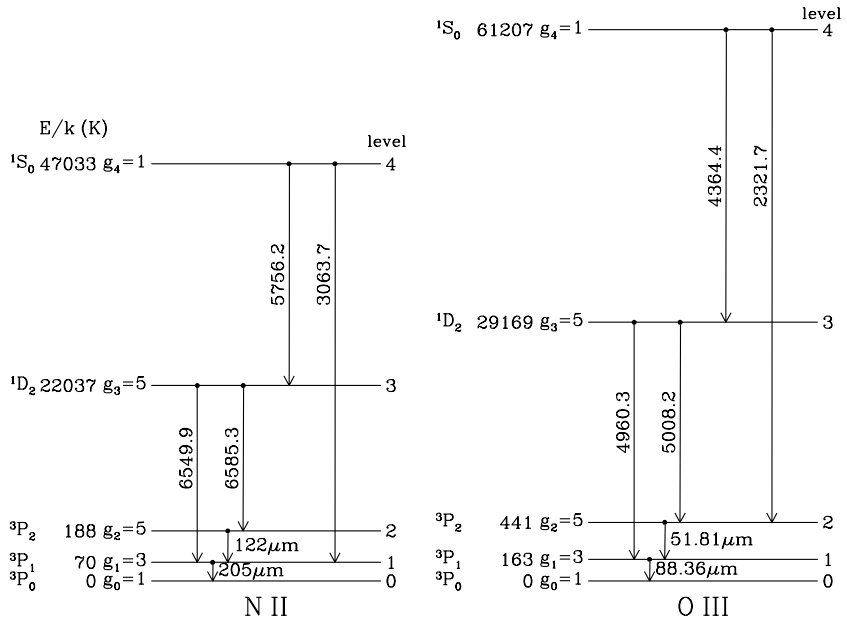


Figure 3.2: *Energy level diagram for the ground state configuration of the $2p^2$ ions* N II *and* O III*. Note that for clarity the fine-structure splitting between $^3P_0$, $^3P_1$, and $^3P_2$ is strongly exaggerated. Forbidden transitions connecting these levels are shown, with wavelengths in vacuo. From: Bruce Draine,* Physics of the interstellar medium and intragalactic medium.

Table 3.1 lists the terms for the ground state configuration of atoms and ions where the outermost subshell is ns or np.

Fine structure

Fine structure is a result of the coupling between the orbital angular momentum **L** of the outer electron and its spin angular momentum **S**. The total angular momentum $\mathbf{J} = \mathbf{L} + \mathbf{S}$ lies between $|L - S| \leq J \leq L + S$.

When $L > 0$ and $S > 0$, there is more than one way to add **L** and **S** to get the total angular momentum $\mathbf{J} = \mathbf{L} + \mathbf{S}$. For given L and S, the allowed values of J range from $|L - S|$ to $L + S$. Each will have different values of $\mathbf{L} \cdot \mathbf{S}$, and will differ in energy due to *electron spin-orbit coupling*. The fractional energy shifts are however small, of order 10^{-2} eV. This splitting of energy levels is referred to as *fine structure* and is also indicated in Fig. 3.2. Beware that in the figure the energy shifts involved are strongly exaggerated. The ^1D and ^1S terms in our example of C I have $S = 0$, therefore do not show fine structure. The ^3P term can have $J = 0, 1, 2$, leading to fine-structure splitting between the three different fine-structure *levels* of the term: $^3P_0$, $^3P_1$, and $^3P_2$. Because of the possibility of multiple J values for a given L

Ground configuration	Terms (ignoring J) and levels (considering J) in order of increasing energy	Examples
... ns^1	$^2S_{1/2}$	H I, He II, C IV, N V, O VI
... ns^2	$^1S_0$	He I, C III, N IV, O V
... np^1	$^2P^\circ_{1/2,3/2}$	C II, N III, O IV
... np^2	$^3P_{0,1,2}$, $^1D_2$, $^1S_0$	C I, N II, O III, Ne V, S III
... np^3	$^4S^\circ_{3/2}$, $^2D^\circ_{3/2,5/2}$, $^2P^\circ_{1/2,3/2}$	N I, O II, Ne IV, S II, Ar IV
... np^4	$^3P_{0,1,2}$, $^1D_2$, $^1S_0$	O I, Ne III, Mg V, Ar III
... np^5	$^2P^\circ_{3/2,1/2}$	Ne II, Na III, Mg IV, Ar IV
... np^6	$^1S_0$	Ne I, Na II, Mg III, Ar III

Table 3.1: *Terms for ns and np subshells. From: Bruce Draine*, Physics of the interstellar medium and intragalactic medium.

and S, the terms are also referred to as *multiplets*. Terms with only one possible value of J are referred to as *singlets*. Terms with two, three and four possible J values are referred to as *doublets*, *triplets*, *quartets*, and so on. The ^3P term is thus a triplet.

The total number of possible quantum states in a term (see also above) is referred to as the *multiplicity*. The multiplicity of a term with total spin S and orbital angular momentum L is

$$g = (2S + 1) \times (2L + 1)\,. \tag{3.9}$$

Thus, the ^3P (i.e., $L = 1$, $S = 1$) term has multiplicity 9. The ^1D (i.e., $L = 2$, $S = 0$) and ^1S (i.e., $L = 0$, $S = 0$) terms have multiplicities 5 and 1. When spin-orbit coupling is taken into consideration, these states are split into distinct fine-structure levels, each with a definite value of J and a *degeneracy*

$$g = (2J + 1)\,. \tag{3.10}$$

For $^3P_0$, $^3P_1$, and $^3P_2$ the values of g are 1, 3, and 5, respectively.

To complete our discussion of the spectroscopic notation, we account for this fine structure, and also introduce the purely quantum mechanical property *parity*. The notation for the angular momentum quantum numbers (in a multi-electron atom) becomes

$$^{2S+1}\mathcal{L}^p_J, \tag{3.11}$$

where $\mathcal{L} = S, P, D, F, ...$ for $L = 0, 1, 2, 3, ...$, and p denotes the parity of the level. It is referred to as the *term symbol*. The parity of an energy level is 'even' or 'odd' depending on whether the electronic wave function changes sign under reflection of all of the electron positions through the origin. Parity is a purely quantum mechanical property. p is left blank for a state of even parity and p is $\circ$ for a state of odd parity. If ℓ_i are the orbital angular momentum quantum numbers of the individual electron orbits, then parity is even if $\Sigma_i\,\ell_i$ is

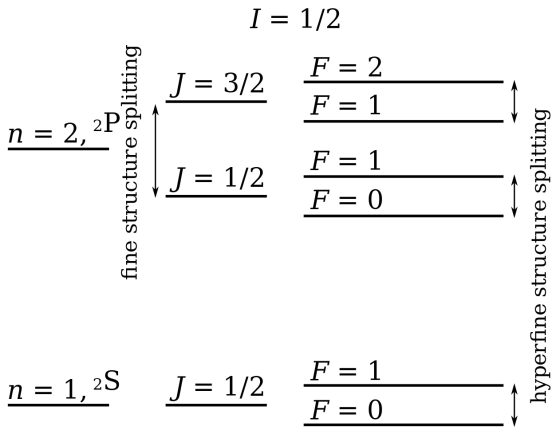


Figure 3.3: *Electronic configuration and energy level splitting. From left to right: an example of an LS term structure, fine structure $\mathbf{J} = \mathbf{L} + \mathbf{S}$, and hyperfine structure $\mathbf{F} = \mathbf{J} + \mathbf{I}$.*

even, and parity is odd if $\Sigma_i \ell_i$ is odd. The ground configuration $1s^2 2s^2 2p^2$ of C I has even parity. Each of the s electrons has orbital quantum number $\ell = 0$ and each of the p electrons has $\ell = 1$. Therefore $\Sigma \ell_i$ is 2, i.e. even.

Hyperfine structure

Hyper fine structure is a result of the coupling of $\mathbf{J}$ with the total nuclear angular momentum $\mathbf{I}_\mathrm{n}$. The total atomic angular momentum $\mathbf{F} = \mathbf{J} + \mathbf{I}_\mathrm{n}$ lies between $|J - I_\mathrm{n}| \leq F \leq J + I_\mathrm{n}$.

So far, we have discussed the state of the electron. We now include in our description properties of the nucleus. If the nucleus has non-zero spin, it will have an intrinsic magnetic moment. In that case, fine structure levels with non-zero electronic angular momentum can themselves be split due to interaction of the electrons with the magnetic field produced by the nucleus. This *hyperfine splitting*[4] is typically of order 10^{-6} eV, corresponding to centimeter wavelengths (see also Fig. 3.3).

In describing both nuclear and electron angular momentum, it is customary to let

$$\begin{aligned} J &\equiv \text{electronic angular momentum} \\ I_\mathrm{n} &\equiv \text{nuclear angular momentum, and} \\ F &\equiv \text{total angular momentum.} \end{aligned}$$

[4]So, fine structure is an interaction between the magnetic field generated by the electron's motion around the nucleus and electron spin, and hyperfine structure is an interaction between the magnetic field generated by the electron's motion around the nucleus and nuclear spin.

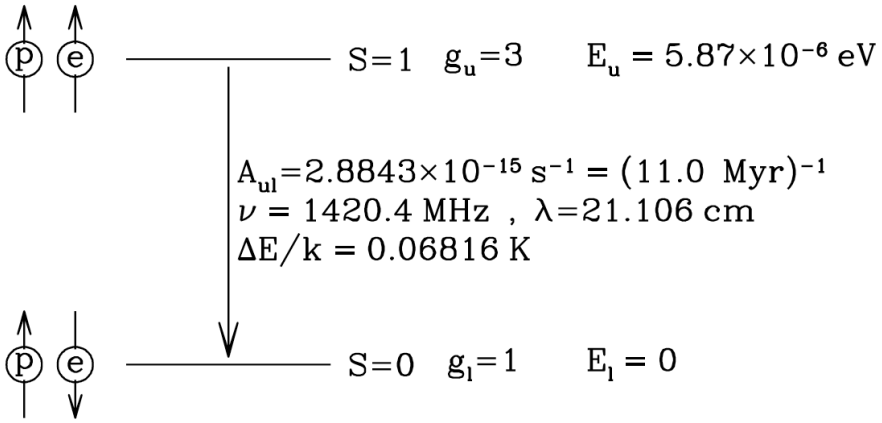


Figure 3.4: *Schematic illustration of hyperfine splitting of the ground state of atomic hydrogen. From: Gould 1994.*

The best-known example of hyperfine splitting is that of the electronic ground state of hydrogen. This ground state has quantum numbers: $n = 1$, $\ell = 0$, $m_\ell = 0$, and $m_s = 1/2$. This $1s\,^2S_{1/2}$ state has electronic angular momentum $J = {}^1\!/_2$. Like the electron, the proton has two spin orientations, denoted $m_i = +{}^1\!/_2$ and $-{}^1\!/_2$, such that $I_\mathrm{n} = {}^1\!/_2$.

In the true ground state of hydrogen, the spins and magnetic moments are anti-parallel, so that the total angular momentum $F = 0$ and the statistical weight $g = 1$. In the exited level, the spins and magnetic moments are parallel, so the total angular momentum $F = 1$ and the statistical weight $g = 3$. See Fig. 3.4 for a graphical representation. The interaction energy due to the magnetic moments is very small

$$\Delta E = 5.9 \times 10^{-6}\ \mathrm{eV} = k \times 0.07\ \mathrm{K} = h \times 1.420\ \mathrm{GHz} = \frac{hc}{21\ \mathrm{cm}}, \tag{3.12}$$

where the last three expressions give the energy in terms of (i) temperature, (ii) frequency, and (iii) wavelength. The spontaneous radiative de-excitation coefficient or Einstein A-coefficient for this transition is $A = 2.87 \times 10^{-15}\,\mathrm{s}^{-1}$. An individual hydrogen atom therefore experiences such a transition typically once per $1/A$ seconds. One finds that a typical hydrogen atom spends on average $\sim 10^7$ years in the upper level before it de-excites spontaneously, with the emission of a 21-cm photon. One might conclude that this radiation would be undetectable. However, the weak emission per hydrogen atom is compensated by the enormous amount of hydrogen in interstellar space. Actually, it is the main signature of cool atomic hydrogen and a very good tracer of the distribution of cold gas in galaxies.

Zeeman effect

In this short discussion on the Zeeman effect, we follow Drain (2011). When a static magnetic field $\mathbf{B}_\circ$ is applied, each of the fine-structure levels $\mathcal{L}_J$ splits into $2J + 1$ energy levels, with

energies depending on the value of $\mathbf{J} \cdot \mathbf{B}_\circ$. The energy splittings are however small, of order $\Delta E = \mu_B B_\circ \sim 5.78 \times 10^{-15} (B_\circ/\mu G)$ eV, where $\mu_B = e\hbar/(2m_e c)$ is the Bohr magneton. Interstellar magnetic field strengths are of order 1–100 μG, and therefore Zeeman shifts are difficult to measure.

Zeeman splitting is most often studied in the hydrogen 21 cm transition, with an energy $\Delta E = 5.9 \times 10^{-6}$ eV. An interstellar magnetic field of 10 μG shifts the frequency in about one part in 10^4, or 3 m/s when expressed in velocity units. This shift is much smaller than the frequency shifts $v/c \sim 10^{-5}$ due to a radial velocity of a few km/s, and it would be nearly impossible to detect, except that it leads to a shift in frequency between the two circular polarization modes. The Zeeman effect in H I λ21cm can therefore be detected by taking the difference of the two circular polarization signals. We will not further pursue this topic in these lecture notes.

3.3 Selection rules

A final and important point to keep in mind when considering transitions between levels is that they obey certain selection rules. Some changes in quantum numbers may occur, others do not – or, at least, are much rarer. The strong transitions in spectra always satisfy what are referred to as the selection rules for electric dipole transitions. These lead to so-called *allowed transitions*. Here, we summarize these rules. We also give the selection rules for *intersystem transitions* and *forbidden transitions* that do not satisfy the electric dipole selection rules but nevertheless are strong enough to be of astrophysical importance.

A complete overview of the selection rules is given in Table 3.2. The table features the total magnetic quantum number M_J which takes the values $M_J = -J, -J+1, ..., J-1, J$. This creates $2J+1$ sublevels or *states* for each value of J. These states are degenerate in the absence of a magnetic field. The splitting of levels into states in a magnetic field is generally known as the *Zeeman effect*.

Intersystem transitions – sometimes also referred to as semi-forbidden or inter-combination transitions – fulfill the electric dipole selection rules except for rule 5 and have $\Delta S \neq 0$. These transitions are considerably weaker than allowed transitions. An intersystem transition is denoted with a single right square bracket – for example N II] λ2143.4Å $^3P_2 - {}^5S^\circ_2$.

A forbidden transition is denoted with two square brackets – for example [N II] λ6549.9Å $^3P_1 - {}^1D_2$. This example fails rule 3 (parity is unchanged) and it fails rule 4 (single electron wave functions are unchanged). Lets clarify the failure to comply with rule 4 a bit more: this is a transition between lower level $1s^2 2s^2 2p^2\, {}^3P_1$ and upper level $1s^2 2s^2 2p^2\, {}^1D_2$. Notice that none of the electrons has changed nl. This is an example of a magnetic dipole transition.

There is a hierarchy in the transition probabilities of the three types of transitions. Very roughly speaking, intersystem transitions are $\sim 10^6$ times weaker than permitted transitions,

		Electric dipole ('allowed')	Electric quadrupole ('forbidden')	Magnetic dipole ('forbidden')
Rigorous rules	1.	$\Delta J = 0, \pm 1$ (not $J = 0 - 0$)	$\Delta J = 0, \pm 1, \pm 2$ (not $J = 0 - 0, \frac{1}{2} - \frac{1}{2}, 0 - 1$)	$\Delta J = 0, \pm 1$ (not $\Delta J = 0 - 0$)
	2.	$\Delta M_J = 0, \pm 1$ (not $M_J = 0 - 0$ if $\Delta J = 0$)	$\Delta M_J = 0, \pm 1, \pm 2$	$\Delta M_J = 0, \pm 1$ (not $M_J = 0 - 0$) if $\Delta J = 0$
	3.	Parity changes	Parity unchanged	Parity unchanged
With negligible configuration interaction	4.	One electron jumping, with $\Delta l \pm 1$, Δn arbitrary	No change in electron configuration; *or* one electron jumping with $\Delta l = 0, \pm 2$, Δn arbitrary	No change in electron configuration; i.e, for all electrons, $\Delta l = 0$, $\Delta n = 0$
For LS coupling only	5.	$\Delta S = 0$	$\Delta S = 0$	$\Delta S = 0$
	6.	$\Delta L = 0, \pm 1$ (not $L = 0 - 0$)	$\Delta L = 0, \pm 1, \pm 2$ (not $\Delta L = 0 - 0, 0 - 1$)	$\Delta L = 0$ $\Delta J = \pm 1$

Table 3.2: *Selection rules for atomic spectra. Rules 1, 2 and 3 must always be obeyed. For electric dipole transitions, intercombination lines violate rule 5 and/or 6. Electric quadrupole and magnetic dipole transitions are also described as forbidden. From: NIST Physical Measurement Laboratory.*

and forbidden transitions are $\sim 10^2 - 10^6$ times weaker than intersystem transitions. Despite begin very 'weak', forbidden transitions do frequently appear in astrophysics because every atom and ion has excited states that can *only* decay via forbidden transitions. At high densities, such excited states would be depolulated by collisions, but at the very low densities of interstellar space, collisions are so rare that there is time for forbidden radiative transitions to take place. We will come back to this point when discussing the two-level system.

As an example we may discuss the energy level diagram of the lowest levels of N II and O III (see Fig. 3.2). It turns out that there are no permitted transitions possible between any of these energy levels: all of these transitions violate selection rule 3. The ^{3}P $\rightarrow$ ^{1}D and ^{3}P $\rightarrow$ ^{1}S transitions in addition violate selection rule 5; the ^{1}D $\rightarrow$ ^{1}S transition also violates rule 6. The fine structure transitions within the ^{3}P level too are forbidden. These all violate rule 5. In addition, the $^3P_0 \rightarrow {}^3P_2$ line violates rule 1. We will see later that the fine structure transitions are quite important for the energy balance of interstellar gas. In addition they are important diagnostic lines of this gas.

Though the example focussed on N II and O II, the exact same arguments holds for all atoms and ions with a similar electron configuration, i.e. all np^2 and np^4 configurations (see Table 3.1).

Exercise 3.1

Classify the following emission lines as either *(i)* permitted, *(ii)* intersystem (a.k.a. intercombination), or *(iii)* forbidden, and give your reason.

(a)	C III	$1s^2 2s2p\ ^3P^\circ_1$	$\rightarrow$	$1s^2 2s^2\ ^1S_0$	1908.7 Å
(b)	O III	$1s^2 2s^2 2p^2\ ^1D_2$	$\rightarrow$	$1s^2 2s^2 2p^2\ ^3P_2$	5008.2 Å
(c)	O III	$1s^2 2s^2 2p^2\ ^1S_0$	$\rightarrow$	$1s^2 2s^2 2p^2\ ^1D_2$	4364.4 Å
(d)	O III	$1s^2 2s2p^3\ ^5S^\circ_2$	$\rightarrow$	$1s^2 2s^2 2p^2\ ^3P_1$	1660.8 Å
(e)	O III	$1s^2 2s^2 2p^2\ ^3P_1$	$\rightarrow$	$1s^2 2s^2 2p^2\ ^3P_0$	88.36 μm
(f)	C IV	$1s^2 2p\ ^2P^\circ_{3/2}$	$\rightarrow$	$1s^2 2s\ ^2S_{1/2}$	1550.8 Å
(g)	Ne II	$1s^2 2s^2 2p^5\ ^2P^\circ_{1/2}$	$\rightarrow$	$1s^2 2s^2 2p^5\ ^2P^\circ_{3/2}$	12.814 μm
(h)	O I	$1s^2 2s^2 2p^3 3s\ ^3S^\circ_1$	$\rightarrow$	$1s^2 2s^2 2p^4\ ^3P_2$	1302.2 Å

Exercise 3.2

Using the Keck telescope at Mauna Kea in Hawai, Antoinette Songaila et al. (1994, Nature 371, 43) have detected the absorption of the lowest fine-structure lines of C I in a cloud at redshift $z = 1.776$, towards the quasar Q1331+170.

a) Give the electron configuration and term symbol of the ground state of C I?

b) The column densities (see Eq. 7.6) that can be derived from the strength of the two absorption lines that feature in this ground level are 7.2×10^{12} cm^{-2} for the $J = 0$ level, and 0.9×10^{12} cm^{-2} for the $J = 1$ level. Compute the excitation temperature $T_{\rm exc}$ that corresponds to the relative occupation of these two levels. The energy of the $J = 1 - 0$ transition is 609.75 μm (or 16.4 cm^{-1}).

c) What does this temperature signify? Is it the kinetic temperature of the gas or the radiation temperature of the diffuse field at $z = 1.776$?

d) If we assume that this is the radiation temperature, what could be the source of the radiation?

4

Molecular structure

In this chapter we discuss the energy-level structure of molecules. In order to understand molecular spectra, we should study the quantum mechanics of molecules. This would lead too far from the topics of this course, and so we restrict ourselves to some basic ideas and a comparison of the classical picture of molecular motions to the quantum mechanical treatment. As in the discussion of atomic levels, here too we make use of the concise summary of Draine (2011), who often relies on the excellent book *Physics of Atoms and Molecules* by B.H. Bransden & C.J. Joachain (2003), to which we refer for a more in-depth treatment of molecular physics. Basics of molecular bondings and spectroscopy can for instance be found in *Physics for Scientists & Engineers, with Modern Physics*, by Giancoli (early editions) or Serway & Jewett (later editions), on which this chapter also relies.

Energy considerations in quantum mechanics start with a formulation of the total energy in the system. This is expressed by the Hamiltonian, which is the sum of the kinetic energies of all the particles plus the electronic potential energy of the particles associated with the system. As a molecule consists of several components (at least two nuclei and one electron), multiple terms will appear in the Hamiltonian and solving for the discrete states that are allowed, subject to a constant total energy E, may quickly lead to involved mathematics as interactions between these components must be considered. Moreover, several types of kinetic energy may need to be considered. In general,

$$E = E_{\rm elec} + E_{\rm trans} + E_{\rm vib} + E_{\rm rot}, \tag{4.1}$$

where $E_{\rm elec}$ denotes electronic potential energy, due to the interactions between the molecule's electrons and nuclei, $E_{\rm trans}$ is translational energy, due to the motion of the molecule's center of mass through space, $E_{\rm vib}$ is vibrational energy, due to the vibration of the molecule's constituent atoms, and $E_{\rm rot}$ is rotational energy, due to the rotation of the molecule about its center of mass. The translational energy is unrelated to internal structure of the molecule, so this molecular energy is unimportant in interpreting molecular spectra. It is in particular the electronic potential energy of a molecule that is very complex, but various techniques have been developed to approximate its values. Although the electronic energies can be studied, significant information about a molecule can be determined by analyzing its quantized rotational and vibrational energy states. These lecture notes focus on the latter two types of

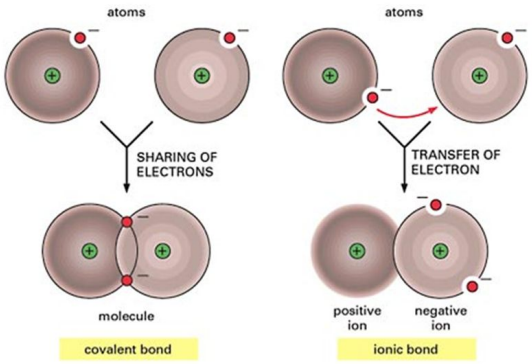


Figure 4.1: *Principle of covalent and ionic bonds. In symmetrical molecules such as H_2, O_2 and Cl_2 the bond is covalent in nature. In a covalent bond the electrons being shared are shared equally among the atoms. In asymmetrical molecules such as NaCl the bond is ionic. In an ionic bond the electrons being shared are shared unequally among the atoms.*

energies as they are most relevant for the medium in-between the stars.

In general one finds that $E_{\rm elec} > E_{\rm vib} > E_{\rm rot}$. Typically, electronic transitions are in the ultraviolet and optical part of the spectrum, vibrational transitions are in the infrared part of the spectrum, and rotational spectra are in the far-IR to radio part of the spectrum.

Before we turn to vibrational and rotational transitions, we first briefly summarize aspects of molecular bonds and the designation of energy levels.

4.1 Molecular bonds

A molecule is a group of two or more atoms that are strongly held together so as to function as a single unit. When atoms attach in such a way we say that a *chemical bond* has been formed. There are two main types of strong chemical bonds: *covalent bonds* and *ionic bonds* (see Fig. 4.1). Many bonds are actually intermediate between these two types.

Covalent and ionic bonds

In a pure covalent bond the electrons are being shared equally among the atoms. Such bonds occur mainly in symmetrical molecules such as H_2, O_2, and Cl_2. Let us take H_2 as an example – the mechanism is basically the same for other covalent bonds. As two H atoms approach each other (be ware that this is *not* a statement on the formation of H_2), the electron clouds begin to overlap, and the electrons from each atom can 'orbit' both nuclei. If both electrons

are in the ground state ($n = 1$) of their respective atoms their spins must be anti-parallel, i.e. $m_s = +1/2$ for the one and $m_s = -1/2$ for the other, so that their total spin $S = 0$. The reason is that because they share the same space, the two electrons can not be in the same state because of the exclusion principle. If the two electrons would have parallel spins (corresponding to $S = 1$) they can not occupy the same space, i.e the electrons spend very little time between the two nuclei. If one would explain this in terms of the electron wave functions, one would state that the exclusion principle requires that when the spins are the same, their is destructive interference of the electron wave functions in the region between the two atoms. But when the spins are opposite, constructive interference occurs in the region between the two atoms, resulting in a large amount of negative charge there. Thus a covalent bond can be said to be the result of constructive interference of the electron wave function in the space between the two atoms, and of the electrostatic attraction of the two positive nuclei for the negative charge concentration between them.

An ionic bond is, in a sense, a special case of the covalent bond. Instead of the electrons being shared equally, they are shared unequally. NaCl is an example of a molecule with an ionic bond. The outer electron of sodium spends most of its time orbiting the chlorine atom, as it feels a much smaller electric attraction to Na than to Cl. Effectively, sodium becomes positively charged (Na^+) and chlorine negatively charged (Cl^-), hence the ionic bond.

As pointed out, a pure covalent bond in which the electrons are shared equally occurs mainly in symmetrical molecules such as H_2. When the atoms involved are different from each other, it is usual to find that the shared electrons are more likely to be in the vicinity of one atom than the other. The extreme case is the ionic bond. In intermediate cases the covalent bond is said to have a *partial ionic character*. Molecules with at least a partial ionic character are *polar* – that is, one part (or parts) of the molecule has a net positive charge and other parts a net negative charge. An example is the water molecule H_2O. The shared electrons are more likely to be found around the oxygen atom than around the two hydrogen atoms. The effect is that there is a net positive charge on each H atom and a net negative charge on the O atom.

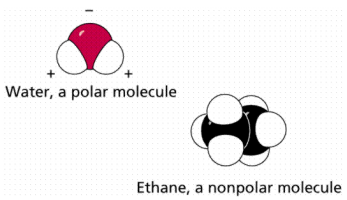


Figure 4.2: *Polar and non-polar molecules.*

Potential-energy diagrams for molecules

The electric forces that are responsible for the bonding in a molecule are related to an electronic potential energy function. A stable molecule is expected at a configuration for which the electronic potential energy function for the molecule has its minimum value. This function should account for two known features of molecular bonding:

- The force is repulsive at very small separation distances. When two atoms are brought close together some of their electron shells start to overlap, resulting in repulsion between the shells partly because of electrostatic forces and partly as a result of the exclusion principle. Some electrons in the overlapping shells are forced into higher energy

states and the system energy increases because of the repulsive force between the nuclei themselves.

- At somewhat larger separations, the force between the atoms is attractive as a result of the covalent and/or ionic bonds.

Accounting for these two features, the electronic potential energy for a system of two atoms can be represented by an expression of the form

$$U(r) = -\frac{A}{r^n} + \frac{B}{r^m}, \tag{4.2}$$

where r is the internuclear separation distance between the two atoms and n and m are small positive integers. The parameter A is associated with the attractive force and B with the repulsive force, as

$$F(r) = -\frac{\mathrm{d}U(r)}{\mathrm{d}r} = -\frac{nA}{r^{n+1}} + \frac{mB}{r^{m+1}}, \tag{4.3}$$

where $F(r)$ is the force. The behavior of $U(r)$ is shown in Fig. 4.3. At large distances between the two atoms, the slope of the curve is positive, corresponding to a net attractive force. At the equilibrium distance $r_\circ$, called the *bond length* or *bond distance*, the attractive and repulsive forces just balance. At this point, the potential energy has its minimum value and the slope of the curve is zero. This minimum potential energy is referred to as the *binding energy*, the *bond energy*, or the *dissociation energy*. For molecular hydrogen, the binding energy is about 4.5 eV and $r_\circ = 0.741\,\text{Å}$. At short distances, the slope is negative, corresponding to a net repulsive force. If only two point sources would be present, $U(r) \propto q_1\, q_2/r$ (where q_i are the charges of the two nuclei), hence $n = 1$.

For many bonds, the potential-energy curve has a shape that is slightly different from that shown in Fig. 4.3: at large distances the potential energy becomes positive. This implies that the atoms do not interact spontaneously. Instead, some additional energy must be injected into the system to get it over the barrier in the potential diagram. This required energy is the *activation energy*. The activation energy often reflects a need to break other bonds, before the one under discussion can be made.

Van der Waals bonds

The binding energies for covalent and ionic bonds is typically 2 to 5 eV. These bonds, which hold atoms together to form molecules, are often called *strong bonds* to distinguish them from so-called *weak bonds*. The term weak bond refers to attachments between molecules due to electrostatic attraction – such as between polar molecules. Binding energies for weak bonds are typically in the range 0.04 to 0.3 eV. Weak bonds are referred to as *van der Waals bonds*, and the forces involved *van der Waals forces*. The potential energy has the general shape shown in Fig. 4.3, with the potential energy varying as $1/r^6$ (i.e. $n = 6$ in Eq. 4.2).

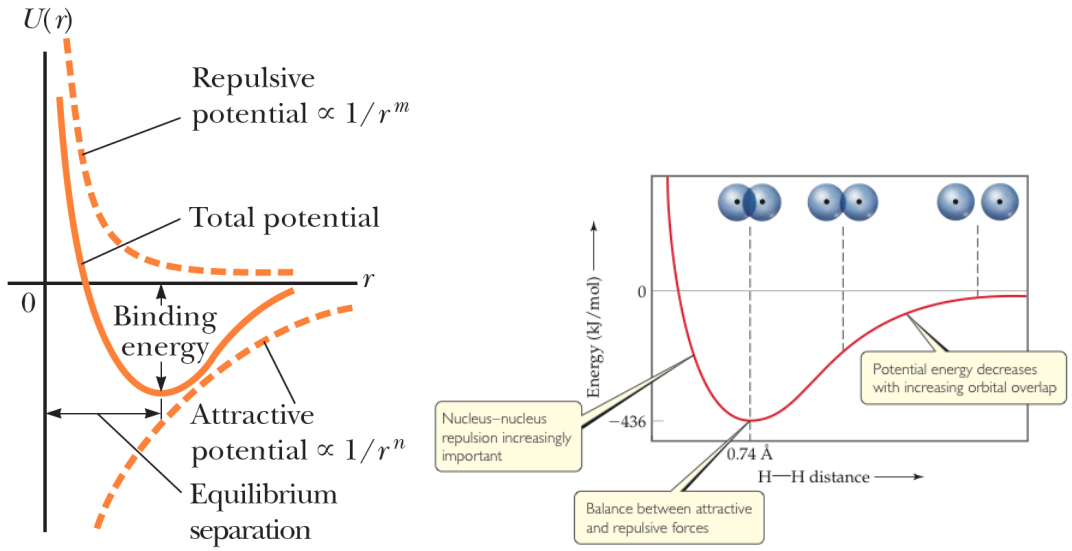


Figure 4.3: *Total potential energy as a function of internuclear separation distance for a system of two atoms. Though the binding energy is indicated to be the energy to the bottom of the potential curve, it is actually the energy to the lowest quantum energy state, which may be slightly above the bottom. See for instance, Fig. 4.10. From: Raymond A. Serway & Jown W. Jewett Jr.,* Physics for Scientists and Engineers with Modern Physics*, 9th edition (left) and Chemistry Stack Exchange (right).*

4.2 Diatomic molecules

The description of molecular structure is considerably more complicated than that of isolated atoms, but fortunately the problem is greatly simplified because the mass of the electron is much smaller than that of the nuclei, while the forces to which the electrons and the nuclei are subjected are of comparable magnitude. As a result, the velocities of the nuclei are much slower than that of the electrons, and the nuclei occupy nearly fixed positions within the molecule. The assumption that the nuclei in the molecule are truly fixed, and that only the electrons are free to move, is known as the *Born-Oppenheimer approximation*. In atoms and ions, the electrons move in a spherically symmetric potential. In molecules, the electrons move in a Coulomb potential due to two or more nuclei, and spherical symmetry does not apply. In the case of *diatomic molecules* – or, more generally, linear molecules – the Coulomb potential due to the nuclei is symmetrical under rotation around the line passing through the two nuclei, the *internuclear axis*.

The projection of the total electronic angular momentum on the internuclear axis is referred to as L_z. It is conventional to define $\Lambda \equiv |L_z|$. Because the potential is axisymmetric, the two states $L_z = \pm\Lambda$ have the same energy. The projection of the total electron spin onto the internuclear axis is S_z. One defines $\Sigma = |S_z|$. The projection of the total electronic angular momentum on the internuclear axis is J_z. If Λ and Σ are both non-zero, then there are two

possible values: $J_z = |\Lambda - \Sigma|$ and $J_z = \Lambda + \Sigma$. States with different $|J_z|$ will differ in energy due to *fine-structure splitting*.

If one or more nuclei have non-zero *nuclear spin* and $J_z \neq 0$, then there will be an interaction between the nuclear magnetic moment and the magnetic field generated by the electrons, leading to *hyperfine splitting*. The energy of these states will depend on the orientation of the nuclear angular momentum (or angular momenta) relative to the internuclear axis. As in atoms and ions, this splitting is small of order $\sim 10^{-6}$ eV.

Designation of energy levels

Diatomic molecules with identical nuclei, such as H_2, N_2 and O_2 are referred to as *homonuclear*. To be clear, the nuclei must be truly identical. HD and $^{16}O^{17}O$ are not homonuclear. The energy levels of homonuclear molecules are designated by term symbols

$$^{(2\Sigma+1)}\mathcal{L}_{u,g}, \tag{4.4}$$

where $\mathcal{L} = \Sigma, \Pi, \Delta, ...$ for Λ = 0, 1, 2, ..., where $\Lambda\hbar$ is the projection of the electron orbital momentum onto the internuclear axis. $\Sigma\hbar$ is the projection of the electron spin angular momentum onto the internuclear axis. In homonuclear molecules, there is an extra symmetry since in addition to the axis of symmetry provided by the internuclear axis, there is a centre of symmetry at the midpoint of the distance between the two nuclei. As a result, the electronic wave functions may split in two sets. Those that remain unaffected by the operation $\mathbf{r_i} \rightarrow -\mathbf{r_i}$ are considered to have an even parity and are denoted by the subscript g and are called *gerade states*. Those that are affected have odd parity and are denoted by the subscript u and are called *ungerade states*.

For the special case of $\mathcal{L} = \Sigma$ states, a superscript $+$ or $-$ is added to the term symbol, i.e.

$$^{(2\Sigma+1)}\Sigma^{+/-}_{u,g}, \tag{4.5}$$

where $+$ is used if the electronic wave function is symmetric under reflection through (all) planes containing the nuclei, and $-$ if the electronic wave function is anti-symmetric under reflection through a plane containing the nuclei. So, a homonuclear diatomic molecule has four non-degenerate Σ states: $\Sigma^+_g, \Sigma^+_u, \Sigma^-_g, \Sigma^-_u$.

In the case of a *heteronuclear* diatomic molecule, e.g. HD, OH, or CO, the energy levels are designated

$$^{(2\Sigma+1)}\mathcal{L}_{J_z}, \tag{4.6}$$

where $\mathcal{L}$ and Σ have the same meaning as for homonuclear diatomic molecules, but now J_z is indicated as a subscript. As for homonuclear molecules, if the term symbol is Σ, then an additional superscript $+/-$ is applied, specifying the symmetry of the wave function under reflection through planes containing the nuclei.

Because a given molecule may have more than one electronic state with the same term symbol, the electronic states are distinguished by a letter X, A, B, ..., a, b, ... appearing in front of the

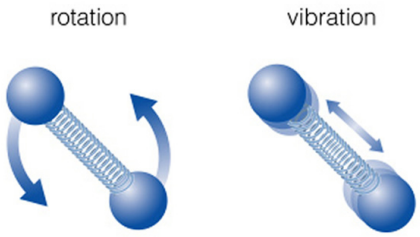


Figure 4.4: *Schematic diagram of rotation and vibration of a diatomic molecule.*

term symbol. The letter X is customarily used to designate the electronic ground state. Examples of electronic ground states are H_2 ($^1\Sigma_g^+$), CH ($^2\Pi_{1/2,3/2}$), CH^+ ($^1\Sigma_0^+$), OH ($^2\Pi_{3/2,1/2}$), CN ($^2\Sigma_{1/2}^+$), CO ($^1\Sigma_0^+$), SiO ($^1\Sigma_0^+$), CS ($^1\Sigma_0^+$). Being the electronic ground state, each species may have an X added before the term symbol. For instance, X $^1\Sigma_g^+$ for H_2.

Vibration and rotation of diatomic molecules

A diatomic molecule can

- vibrate (stretch) along the internuclear axis, and
- rotate around an axis perpendicular to the internuclear axis.

These angular momentum modes add (vectorially) to the electronic angular momentum. The ro-vibrational levels of diatomic molecules are specified by a single vibrational quantum number v and rotational quantum number J. Transitions will change J by either 0, ± 1, or ± 2. It is custom to identify transitions by specifying the upper and lower electronic states, upper and lower vibrational states, and a letter code, O, P, Q, R, S, of which the usages is given in Tab. 4.1. Thus, for example, a transition from the $v = 0, J = 1$ level of the ground electronic state to the $v = 5, J = 2$ level of the first electronic excited state would be written A-X 5-0 $R(1)$. So, in the notation the upper level precedes the lower level.

The H_2 molecule

The electronic ground state of H_2 has zero electronic orbital angular momentum $L = 0$, has zero electron spin $S = 0$, is symmetric under reflection through the center of mass (g), and is symmetric under reflection through planes containing the nuclei ($+$). The ground state is X $^1\Sigma_g^+$.

In the case of H_2, the electronic wave function is required to be anti-symmetric under exchange of the two electrons. The two protons, just like the electrons, are identical fermions, and

Designation	$(J_u - J_l)$	Note
$O(J_l)$	-2	Electronic quadrupole transition
$P(J_l)$	-1	Electronic dipole transition
$Q(J_l)$	0	Electric dipole or electric quadrupole, $Q(0)$ is forbidden
$R(J_l)$	+1	Electric dipole transition
$S(J_l)$	+2	Electric quadrupole transition

Table 4.1: *Usage of O, P, Q, R, S. The lower state is denoted by l, the upper state by u.*

therefore the Pauli exclusion principle anti-symmetry requirement also applies to exchange of two protons. The protons are spin $^1/_2$ particles, therefore the two protons together can have total spin 1 (if the spins are aligned to each other, i.e. parallel) or total spin 0 (anti-parallel). Without going into the quantum mechanics, the consequence of the anti-symmetry requirement is that if the protons have spin 0, the rotational quantum number J must be even. This is referred to as *para* H_2, with $J = 0, 2, 4, ...$ If the two protons are parallel, with total spin 1, the rotational quantum number J must be odd. This is referred to as *ortho-*H_2, with $J = 1, 3, 5, ...$ Because the nuclear spins are only weakly coupled to the electromagnetic field, ortho-H_2 and para-H_2 behave as almost distinct species, with conversion of ortho to para, or vice versa, happening only very slowly at low temperatures (in the absence of a catalyst). See Fig. 4.6 for the ro-virational energy level diagram of H_2.

Because H_2 consists of two atoms of identical mass, the center of mass and the center of charge coincide. Consequently, it has no permanent dipole moment and emits no electric dipole radiation (characterized by $\Delta J \pm 1$). The vibrational states and the rotational states radiate very weakly, via the time-variation of the electric quadrupole moment as the molecule vibrates or rotates. Because the nuclear spin state does not change, the ro-vibrational radiative transitions of H_2 must have $\Delta J = 0$ or $\Delta J = \pm 2$, i.e. ortho to ortho or para to para. Conversion of para-H_2 into ortho-H_2, or vice versa, occurs, as mentioned, only very slowly.

4.3 Rotational spectra of diatomic molecules

It is instructive to use classical mechanics when considering the spectra of molecules. In this case the rotation energy is given by

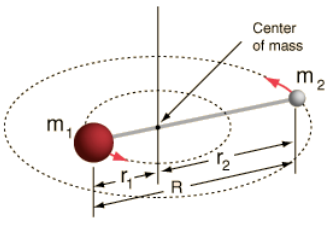


Figure 4.5: *Center of Mass.*

$$E_{\rm rot} = \frac{1}{2} I \omega^2, \tag{4.7}$$

where ω is the angular velocity, i.e. $\omega = 2\pi \nu_{\rm rot}$, where $\nu_{\rm rot}$ is

the rotational frequency, and

$$I = \sum_i m_i \, r_i^2 \tag{4.8}$$

is the moment of inertia. Each of the components i of the system has a mass m_i and a distance r_i to the rotation axis. For the center of mass of a diatomic molecule it holds that $m_1 r_1 = m_2 r_2$, where you should note the distances r_i are absolute distances. Denoting the absolute separation between the two nuclei or internuclear separation as $r = r_1 + r_2$ (see Fig. 4.5), we can express the moment of inertia as follows:

$$\begin{aligned} I &= m_1 r_1^2 + m_2 r_2^2 \\ r_1 &= \frac{m_2}{m_1 + m_2} r \\ r_2 &= \frac{m_1}{m_1 + m_2} r \\ \Rightarrow I &= \frac{m_1 m_2}{m_1 + m_2} r^2 \equiv \mu r^2 \end{aligned} \tag{4.9}$$

So we find that the moment of inertia of a molecule is proportional to the reduced mass μ of the system. Expressing the rotational energy in the angular momentum of the rotating molecule $L = I\omega$ yields

$$E_{\rm rot} = \frac{L^2}{2I}, \tag{4.10}$$

In the quantum-mechanical context the reduced mass μ also plays a role. Moreover, the angular momentum is quantized and given by $L^2 = \ell(\ell+1)\hbar^2$, in which ℓ is zero or a positive integer. The rotational energy of the molecule becomes

$$E_{\rm rot} = \frac{\ell(\ell+1)\hbar^2}{2I}. \tag{4.11}$$

Note that for the ground state $E_{\rm rot} = 0$. If we would do a proper quantum mechanical treatment, and write down and solve the Schrödinger equation for rotation, we would have obtained

$$E_{\rm rot} = \frac{J(J+1)\,\hbar^2}{2I} = \frac{J(J+1)\,h^2}{8\pi^2 \mu r^2} = \frac{J(J+1)\,h^2}{8\pi^2 I} = hc\,BJ(J+1). \tag{4.12}$$

Here, J is the rotational quantum number (not to be confused with the J quantum number for atoms/ions). The constant

$$B = \frac{h}{8\pi^2 c\, I} = \frac{\hbar}{4\pi c\, I} \tag{4.13}$$

is the *rotational constant*. The dimension of B is cm^{-1}. Let us compute B for the two important molecules H_2 and CO. For H_2, the reduced mass $\mu = 0.5\, m_{\rm H}$ and the internuclear separation $r = 0.742\,\text{Å}$, such that the moment of intertia $I_{H_2} = 4.61 \times 10^{-41}\,\mathrm{gr\,cm^2}$. This results in $B = 60.80\,\mathrm{cm}^{-1}$. For CO, $\mu = 6.857\, m_{\rm H}$ and $r = 1.1283\,\text{Å}$, implying $I_{\rm CO} = 1.45 \times 10^{-39}\,\mathrm{gr\,cm^2}$. This results in $B = 1.922529\,\mathrm{cm}^{-1}$.

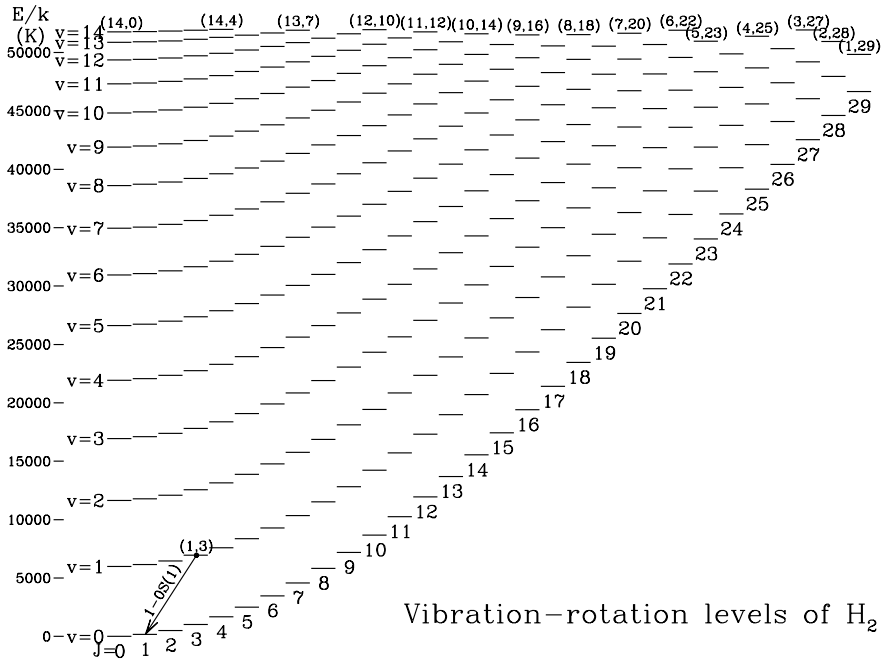


Figure 4.6: *Vibration-rotation energy levels of the ground electronic state of H_2 with $J \leq 29$. The $(v, J) = (1, 3)$ level and the 1-0 S(1) transition at 2.1218 μm are indicated. Transitions between even J values belong to para H_2, those between odd J values to ortho H_2. From: Bruce Draine,* Physics of the interstellar medium and intergalactic medium.

As in atoms, also for molecules we can derive *selection rules* for rotational spectra. These state that for electric dipole radiation

$$\Delta J = \pm 1. \tag{4.14}$$

Therefore the energy difference between the transition J to $J - 1$ is given by

$$E_{\text{photon}} = \Delta E_{\text{rot}} = E(J) - E(J-1) = 2hc\,BJ. \tag{4.15}$$

With this information we can now easily calculate the pure rotational spectrum of a simple rigidly rotating diatomic molecule, provided we know the value of the rotational constant.

For the rigid rotator the spectrum thus consists of a set of evenly spaced lines in frequency space. If centrifugal distortion is considered, it will destroy this constant separation (see next page). Table 4.2 summarises the characteristics of rotational transitions of several diatomic molecules (among others). The $J = 1 - 0$ transitions of molecules consisting of heavy atoms fall at a few millimeters. Because of the lower mass (therefore lower reduced mass μ, hence larger $B \propto 1/\mu$), the rotational energy levels of hydrides (i.e. diatomic molecules in which one is H) are shifted to higher energies, i.e. the sub-millimeter or far-infrared range.

Species	Transitions	ν_{lu} (GHz)	$\lambda_{\rm lu}$ (μm)	E_u (K)	A_{ul} (s^{-1})	$n_{\rm cr}$ cm^{-3}
CO	$1-0$	115.3	2600.1	5.5	7.2×10^{-8}	1.1×10^{3}
	$2-1$	230.8	1298.9	16.6	6.9×10^{-7}	6.7×10^{3}
	$3-2$	346.0	866.4	33.2	2.5×10^{-6}	2.1×10^{4}
	$4-3$	461.5	649.6	55.4	6.1×10^{-6}	4.4×10^{4}
	$5-4$	576.9	519.7	83.0	1.2×10^{-5}	7.8×10^{4}
	$6-5$	691.2	433.7	116.3	2.1×10^{-5}	1.3×10^{5}
	$7-6$	806.5	371.7	155.0	3.4×10^{-5}	2.0×10^{5}
CS	$1-0$	49.0	6118.2	2.4	1.8×10^{-6}	4.6×10^{4}
	$2-1$	98.0	3059.1	7.1	1.7×10^{-5}	3.0×10^{5}
	$3-2$	147.0	2039.4	14.0	6.6×10^{-5}	1.3×10^{6}
	$5-4$	244.9	1224.1	35.0	3.1×10^{-4}	8.8×10^{6}
	$7-6$	342.9	874.3	66.0	1.0×10^{-3}	2.8×10^{7}
	$10-9$	489.8	612.1	129.0	2.6×10^{-3}	1.2×10^{8}
HCO^+	$1-0$	89.2	3360.9	4.3	3.0×10^{-5}	1.7×10^{5}
	$3-2$	267.6	1120.3	26.0	1.0×10^{-3}	4.2×10^{6}
	$4-3$	356.7	840.46	43.0	2.5×10^{-3}	9.7×10^{6}
HCN	$1-0$	88.6	3383.7	4.3	2.4×10^{-5}	2.6×10^{6}
	$3-2$	265.9	1127.5	26.0	8.4×10^{-4}	7.8×10^{7}
	$4-3$	354.5	845.7	43.0	2.1×10^{-3}	1.5×10^{8}
H_2CO	$2_{12}-1_{11}$	140.8	2129.2	6.8	5.4×10^{-5}	1.1×10^{6}
	$3_{13}-2_{12}$	211.2	1419.5	17	2.3×10^{-4}	5.6×10^{6}
	$4_{14}-3_{13}$	281.5	1065.0	30	6.0×10^{-4}	9.7×10^{6}
	$5_{15}-4_{14}$	351.8	852.2	47	1.2×10^{-3}	2.6×10^{7}
NH_3	(1,1) inversion	23.7	12649.5	1.1	1.7×10^{-7}	1.8×10^{3}
	(2,2) inversion	23.7	12649.5	42	2.3×10^{-7}	2.1×10^{3}
H_2	$2-0$	1.06×10^{4}	28.3	510	2.9×10^{-11}	10
	$3-1$	1.76×10^{5}	17.0	1015	4.8×10^{-10}	300

Table 4.2: *Characteristics of rotational molecular cooling lines.* $E_u = h\nu_{lu}/k$. *Critical densities for CO are approximated from Fig. 10 in Yang et al. 2010 (ApJ 718, 1062) at a kinetic temperature of 10 K. From: Tielens (2005), The Physics and Chemistry of the Interstellar Medium.*

Not all molecules show rotational spectra through permitted dipole radiation. Only molecules that have a permanent dipole moment emit in this way. We can understand this by considering that a radiative transition is the result of an accelerated charge. A molecule without permanent dipole moment will not show a change in the spatial distribution of charge when it rotates, and so no net acceleration of charges. No permitted dipole radiation is thus expected. Examples are H_2, C_2, O_2, CH_4. These can still be studied in electric quadrupole radiation with $\Delta J = 2$, but the transition probabilities for these transitions are orders of magnitude smaller than those of permitted ones. The implication is that the most abundant molecule in space, H_2, is very difficult to detect, especially when the gas is cold. Of course there are permitted *vibrational* transitions of H_2 available (see below) but these are often not in thermal equilibrium, are associated with mostly hot molecular gas, and not easy to use as diagnostic probes.

As pointed out above, the energy levels of the quadrupole radiation of H_2 divide out into two

separate ladders with even or odd J. These ladders are not connected by radiative transitions. Because of symmery considerations, the even states correspond to para-H_2 (i.e. anti-parallel nuclear spins) and the odd levels to ortho-H_2 (e.g., parallel nuclear spins). Owing to nuclear spin statistics ($g_S = 2S + 1$), the ortho levels have three times the statistical weight of the para levels (see Tielens 2005).

The lines in a pure rotational spectrum are not equally spaced

Experimental verification of the pure rotational spectrum of molecules show that the lines are *not* equally spaced as the rigid rotator approximation predicts. The discrepancy can be resolved by realizing that a chemical bond is not truly rigid. As the molecules rotate more energetically (which is the case for increasing J), the centrifugal force causes the bond to stretch slightly – resulting in a larger moment of inertia. This small effect can be treated by perturbation theory, and the end result is that the energy can be written as

$$E(J) = hcBJ(J+1) - hcDJ^2(J+1)^2, \tag{4.16}$$

where D is the centrifugal distortion constant. Of course, $D \ll B$. For CO, $D = 6.1193 \times 10^{-6}$ cm^{-1}. The more precise energy difference between the transition J to $J-1$ (compared to Eq. 4.15) is given by

$$E_{\rm photon} = \Delta E_{\rm rot} = E(J) - E(J-1) = 2hc\,BJ - 4hcDJ^3. \tag{4.17}$$

4.4 Cold gas temperature diagnostics: rotational diagram

A straightforward application of pure rotational line emission is to measure column densities (see Eq. 4.23) for different levels of a particular molecule, and to combine these into a *rotational diagram* which can be used to derive the *excitation temperature* of the molecules. We start out with considering the number density, which is related to the excitation temperature through Boltzmann equation (2.70),

$$n_J \propto g_J \, \exp\left[-E(J)/kT\right] = (2J+1) \, \exp\left[-E(J)/kT\right], \tag{4.18}$$

where g_J is the statistical weight of the rotational level. We may substitute $E(J)$ by Eq. (4.12), which yields

$$n_J \propto (2J+1) \exp\left[-BJ(J+1)\, hc/kT\right]. \tag{4.19}$$

This means that the maximum occupation is not always at $J = 0$, but can be at higher levels depending on the temperature of the gas (see Fig, 4.7). One may show that the J for which the population reaches a maximum, $J_{\rm max}$, is reached for

$$J_{\rm max} = \sqrt{\frac{kT}{2hcB}} - \frac{1}{2} = 0.5896 \sqrt{\frac{T}{B}} - \frac{1}{2}. \tag{4.20}$$

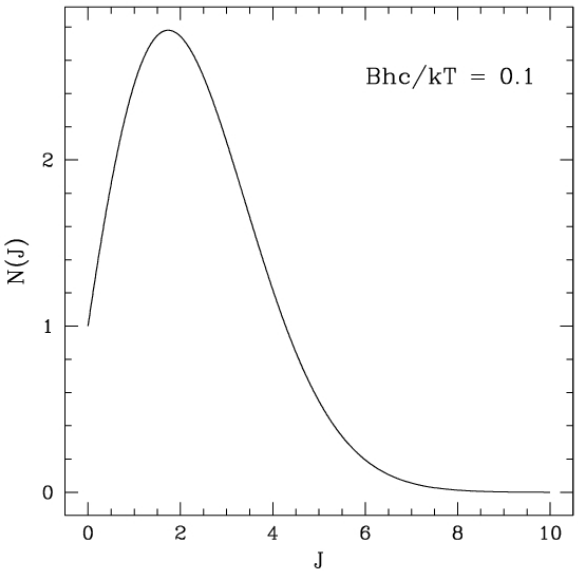


Figure 4.7: Occupation of rotational energy levels for a Boltzmann distribution of molecules for a given kinetic temperature of the gas.

We now turn to the construction of rotational diagrams. Consider the example of a molecular cloud in which the pure rotational lines of some molecule have been measured. We assume all these lines to be optically thin. The energy emitted by the cloud per second in the transition $J \rightarrow J-1$ is

$$L_{\rm cloud} = \int_{\rm Volume} \frac{hc}{\lambda}\, n_J\, A_{J,J-1}\, dV, \tag{4.21}$$

where we need to integrate over the volume of the cloud. Each second $n_J\, A_{J,J-1}$ de-excitations take place in each cubic centimeter of the cloud, where $A_{J,J-1}$ is the Einstein A coefficient for the transition (see e.g. Table 4.2). By multiplying by the energy $h\nu = hc/\lambda$, and integrating over the entire volume, we get the line luminosity $L_{\rm cloud}$ of the cloud integrated over the spectral line profile. We assume that the cloud is spatially resolved, such that we can measure the total (or frequency integrated) line intensity at some spot in our detector. This total intensity is

$$I(J, J-1) = \int_{\rm line\,of\,sight} \frac{hc}{\lambda} n_J \frac{1}{4\pi} A_{J,J-1}\, ds = \frac{hc}{\lambda} \frac{N_J\, A_{J,J-1}}{4\pi}, \tag{4.22}$$

where N_J is the *column density* along the path through the cloud on which we focus, i.e.

$$N_J = \int_{\rm line\,of\,sight} n_J\, ds. \tag{4.23}$$

The factor 4π enters because the total intensity is defined per unit solid angle (the total solid angle being 4π). For the column density we thus find that

$$N_J = \frac{4\pi\, I(J, J-1)}{A_{J,J-1}} \frac{\lambda}{hc} \tag{4.24}$$

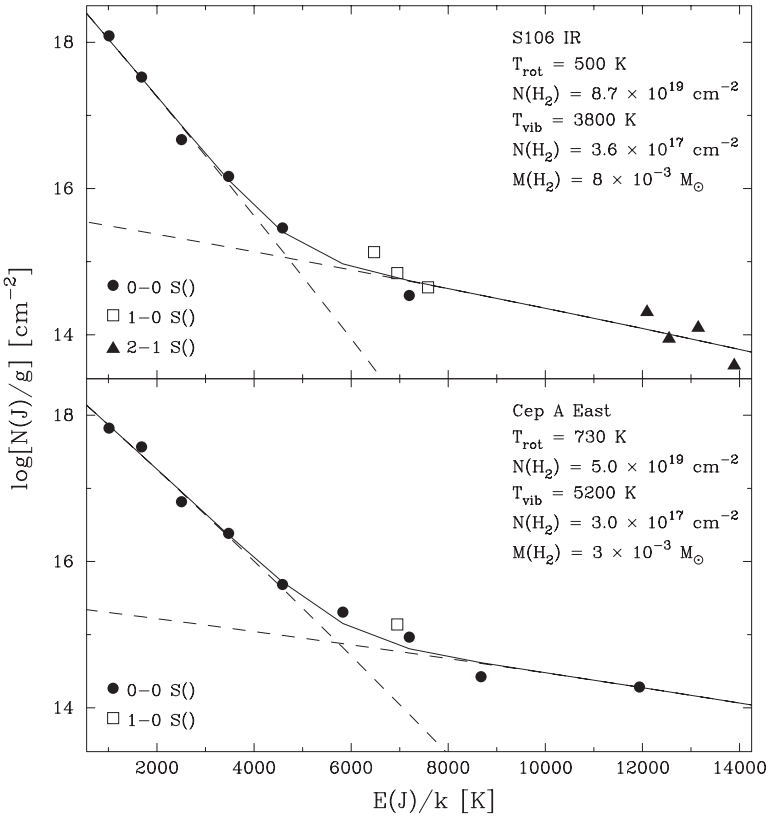


Figure 4.8: H_2 excitation diagram for S106 (top) and Cep A East (bottom). Circles indicate ISO observations of quadrupole transition pure rotational lines. The triangles and squares indicate ground-based measurements of ro-vibrational lines. The dashed lines give Boltzmann distribution fits to the low-lying pure rotational lines and the lines with upper level energies above 5000 K. The solid line is the sum of both contributions. Figure taken from van den Ancker et al., 2000 A&A 358, 1035.

Using Eq. (4.18) it is easy to see that if $\ln(N_J/g)$ is plotted versus $E(J)/k$ the slope of this curve gives the excitation temperature of the molecules averaged (in some weighted way) along the line-of-sight. If the molecules are in LTE, this excitation temperature is the same as the kinetic temperature. We show examples of such rotational diagrams in Fig. 4.8 for two massive young stars embedded in their parental molecular cloud. The pure rotational lines of H_2 have been measured using the *Infrared Space Observatory* (ISO) satellite. Two linear fits are needed to describe the observations. The lower temperature is probably related to the actual kinetic temperature of the H_2 in the cloud. The higher temperature, which is determined by rotational lines of vibrationally exited levels, probably does not reflect the kinetic temperature of the gas but a radiation temperature (implying these transitions are in

non-LTE)

4.5 Vibrational spectra of diatomic molecules

In addition to rotational motion around their center of mass, molecules also experience radial, or vibrational motion. For the case of a diatomic molecule the two nuclei vibrate along the internuclear axis, as shown in Figs. 4.4 and 4.9. As in the case for rotational spectra, it is insightful to consider the classical analogue of the motion, which is the harmonic oscillator. As long as the atoms in the molecule are not too far from their equilibrium positions, the potential energy function is in good approximation parabolic, varying as the square of the position of the particle relative to the equilibrium position $r_\circ$. Figure 4.9 depicts the situation: for separations close to $r_\circ$, the shape of the potential energy curve closely resembles the parabolic shape of the potential energy function of the simple harmonic motion model. For the equation of motion of the harmonic oscillator, it follows that

$$F = -k\,x = \mu \frac{d^2x}{dt^2}, \tag{4.25}$$

where F is the restoring force, $x = r - r_\circ$ is the displacement from the equilibrium position, k is the force constant (or spring constant) and μ the reduced mass. This is *Hooke's law for a restoring force*. Because $F = -dV/dx$, where V is the potential, it follows that

$$V = \frac{1}{2}kx^2. \tag{4.26}$$

The solution to Eq. (4.25) is

$$x = x_0 \sin(2\pi\,\nu_{\rm osc}\,t + \phi), \tag{4.27}$$

where ϕ is the phase determining the starting point of the sine wave and

$$\nu_{\rm osc} = \frac{1}{2\pi}\sqrt{\frac{k}{\mu}} \tag{4.28}$$

is the fundamental or *harmonic frequency* of vibration.

The quantum mechanical equivalent of the classical oscillator is

$$\frac{d^2\psi}{dx^2} + \frac{8\pi^2\mu}{h^2}\left(E - \frac{1}{2}kx^2\right)\psi = 0, \tag{4.29}$$

where ψ is the wavefunction. The possible solutions of this Schrödinger equation are

$$\begin{aligned} E(v) &= \frac{h}{2\pi}\sqrt{\frac{k}{\mu}}\left(v + \frac{1}{2}\right) \\ &= h\nu_{\rm osc}\left(v + \frac{1}{2}\right), \end{aligned} \tag{4.30}$$

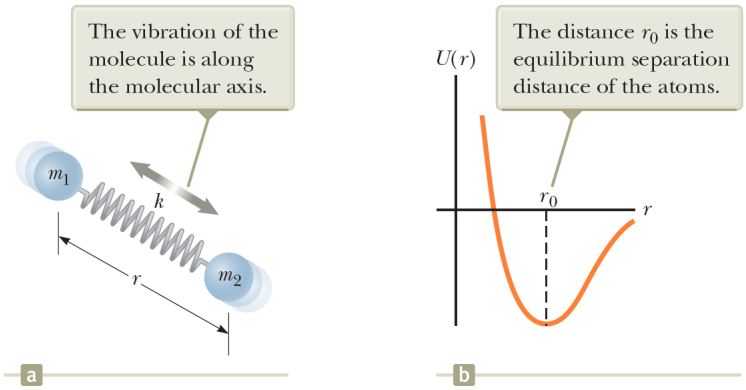


Figure 4.9: *Effective-spring model of a diatomic molecule and plot of the potential energy of a diatomic molecule versus atomic separation distance. For separations close to $r_\circ$, the shape of the potential energy curve closely resembles a parabolic shape. From: Raymond A. Serway & Jown W. Jewett Jr.,* Physics for Scientists and Engineers with Modern Physics, *9th edition.*

which introduces the *vibrational quantum number* $v = 0, 1, 2, ...$ The energies are quantized, which is a general feature of quantum-mechanical systems when a particle is confined. These discrete energy levels are equally spaced and the ground state is not at zero energy – as in the case of rotational spectra – but at an energy $\frac{1}{2}h\nu_{\text{osc}}$. Vibrational transitions are also subject to selection rules. Pure vibrational modes have

$$\Delta v = \pm 1, \tag{4.31}$$

therefore

$$E_{\text{photon}} = \Delta E_{\text{vib}} = \frac{h}{2\pi}\sqrt{\frac{k}{\mu}} = h\nu_{\text{osc}}. \tag{4.32}$$

This implies that for pure vibrational modes there is only one line in the vibrational spectrum.

At most temperatures that we will encounter the molecules will have vibrational energies corresponding to the $v = 0$ state because the spacing between the vibrational energy states is much greater than kT, where k is Boltzmann's constant and T is the temperature. To give an example, the wavelength of the photon that causes the $v = 0 - 1$ transition in the CO molecule is 4.67 μm or 6.42×10^{13} Hz, corresponding to a temperature of 3081 K.

The an-harmonic case and overtones in vibrational spectra

Fig. 4.10 shows that in reality the exact shape of the internuclear potential energy is not a simple parabola. One may improve on the description of the potential by expanding $V(r)$ in

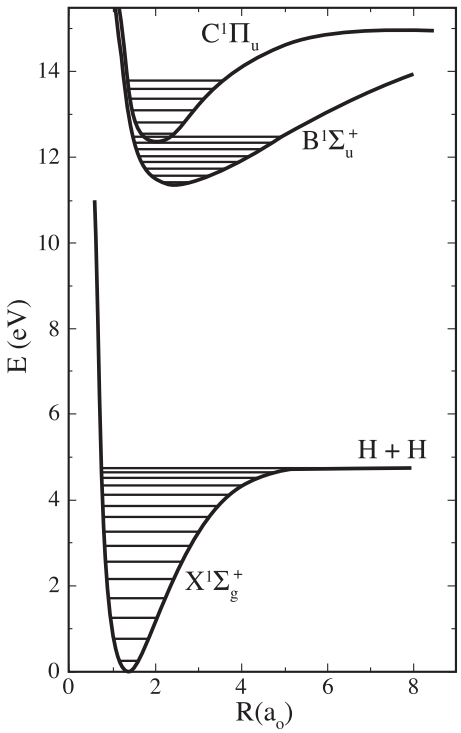


Figure 4.10: *Effective internuclear potential for H_2 for the ground state $X\,^1\Sigma_g^+$ and the first two electronic excited states $B\,^1\Sigma_u^+$ and $C\,^1\Pi_u$. The horizontal lines denoted the vibrational level, with $v = 0$ the ground vibrational level, being the lowest horizontal line (note it is not at zero energy). The x-axis shows the internuclear distance. From: Bruce Draine,* Physics of the interstellar medium and intragalactic medium.

a Taylor series about $r_\circ$, to give

$$\begin{aligned} V(r) - V(r_\circ) &= \frac{1}{2!}\left(\frac{\mathrm{d}^2V}{\mathrm{d}r^2}\right)_{r_\circ}(r-r_\circ)^2 + \frac{1}{3!}\left(\frac{\mathrm{d}^3V}{\mathrm{d}r^3}\right)_{r_\circ}(r-r_\circ)^3 + \cdots \\ &= \frac{1}{2}kx^2 + \frac{1}{6}\gamma x^3 + \cdots \end{aligned} \tag{4.33}$$

The harmonic-oscillator approximation consist of keeping only the quadratic term in Eq. 4.33, and it predicts a single line in the vibrational spectrum. Experimental data show there is, indeed, one dominant line (called the *fundamental*) but also lines of weaker intensity at almost integral multiples of the fundamental. These lines are called *overtones*. If the an-harmonic terms in Eq. (4.33) are included, the Schrödinger equation can be solved by perturbation theory to give

$$E(v) = h\nu_{\mathrm{osc}}\left(v+\frac{1}{2}\right) - \tilde{x}h\nu_{\mathrm{osc}}\left(v+\frac{1}{2}\right)^2 + \cdots, \tag{4.34}$$

where $\tilde{x}$ is the an-harmonicity constant. The an-harmonic correction is much smaller than the harmonic term, i.e. $\tilde{x} \ll 1$. The levels are not equally spaced as they are for a harmonic oscillator; their separation decreases with increasing v (see Fig. 4.11). The harmonic-oscillator approximation is best for small values of v, which are the most important values at the temperatures in interstellar space.

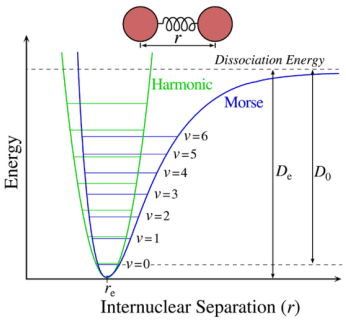


Figure 4.11: *Harmonic (parabolic) potential and an-harmonic (Morse) potential. From: Mark Somoza.*

The selection rule for an an-harmonic oscillator is that Δv can have any integral value, although the intensities of the $\Delta v = \pm 2, \pm 3, ...$ transitions are much less than for the $\Delta v \pm 1$ transitions. The photon energies of $v \rightarrow v'$ transitions are given by

$$\begin{aligned} E_{\rm photon} &= E(v) - E(v') = h\nu_{\rm osc}\left(v - v'\right) \\ &\quad -\tilde{x}h\nu_{\rm osc}\left[v(v+1) - v'(v'-1)\right] \end{aligned} \tag{4.35}$$

for $v = 1, 2, ...$ and $v > v'$. The fundamental frequency ($\Delta v = 1$, e.g., $v = 1 \leftrightarrow 0$ or $v = 2 \leftrightarrow 1$) is less than the frequency for pure harmonic motion. The overtones ($\Delta v = 2$, e.g., $v = 2 \leftrightarrow 0$ or $v = 3 \leftrightarrow 1$ is the first overtone; $\Delta v = 3$, e.g., $v = 3 \leftrightarrow 0$ or $v = 4 \leftrightarrow 1$ the second overtone, etcetera) are not quite integral multiples of the fundamental frequency.

Ro-vibrational transitions

In general, a molecule vibrates and rotates simultaneously. To a first approximation, these motions may be regarded independent of each other, so that the total energy of the molecules for these motions is the sum of Equations (4.12) and (4.30), i.e.

$$E(v, J) = h\nu_{\rm osc}\left(v + \frac{1}{2}\right) + hc\,BJ\,(J+1)\,. \tag{4.36}$$

For each allowed value of the vibrational quantum number v, there is a complete set of rotational levels corresponding to $J = 0, 1, 2, ...$ When a molecule absorbs a photon with the appropriate energy, the vibrational quantum number v increases by one unit while the rotational quantum number J either increases or decreases by one unit (see Fig. 4.12). The molecular absorption spectrum thus consists of two groups of lines: one group to the right of center (in the right panel of Fig. 4.12) and satisfying the selection rules $\Delta J = +1$ and $\Delta v = +1$, and the other group to the left of center and satisfying the selection rules $\Delta J = -1$ and $\Delta v = +1$. The first series is referred to as the P-branch, the second series is the R-branch. The energies of the absorbed photons are

$$E_{\rm photon} = \Delta E = h\nu_{\rm osc} + 2hc\,B(J+1) \qquad J = 0, 1, 2, ... \quad {\rm P(J) - branch} \tag{4.37}$$

$$= h\nu_{\rm osc} - 2hc\,BJ \qquad J = 1, 2, 3, ... \quad {\rm R(J) - branch} \tag{4.38}$$

In certain conditions the transition $\Delta J = 0$ is permitted as well. This means that a third branch will be visible in the spectrum, the so-called Q-branch. (Note that also an O-branch ($\Delta J = -2$) and S-branch ($\Delta J = +2$) exist).

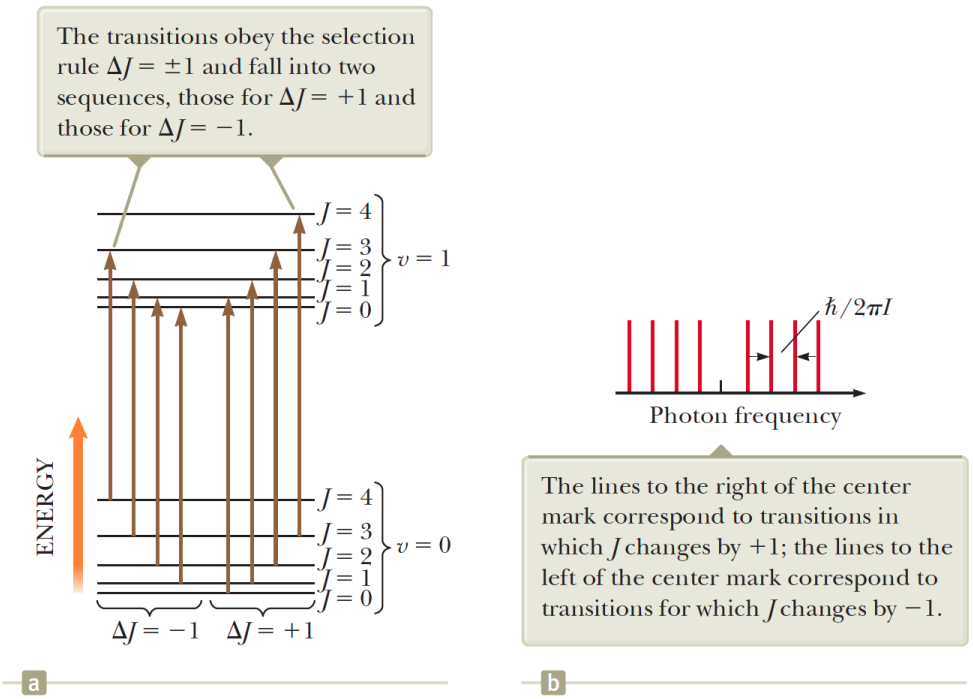


Figure 4.12: *Left: Absorptive transitions between the* $v = 0$ *and* $v = 1$ *vibrational state of a diatomic molecule. Right: expected lines in the absorption spectrum of a molecule. The lines that correspond to* $\Delta J = J_{\text{lower}} - J_{\text{upper}} = +1$ *changes are said to be part of the P-branch; those that correspond to* $\Delta J = -1$ *belong to the R-branch. Recall that* $E_{\text{photon}} = 2hcBJ$*, see Eq. 4.15, therefore* $\nu = \hbar J/2\pi I$*. From: Raymond A. Serway & Jown W. Jewett Jr.,* Physics for Scientists and Engineers with Modern Physics, *9th edition.*

The experimental emission spectrum of the HCl molecule is shown in Fig. 4.13. Notice that each line is split into a doublet. This doubling occurs because two chlorine isotopes (^{35}Cl and ^{37}Cl) were present in the sample to obtain this spectrum. Because the isotopes have different masses, the two HCl molecules have different values of moment of inertia I.

The specific intensity of the lines is proportional to the number of molecules in each of the excited J states (see Eq. 4.19), therefore

$$I \propto n(J) \propto (2J+1) \exp\left[-BJ(J+1)\, hc/kT\right], \tag{4.39}$$

where I is the specific intensity (do not get confused with the moment of inertia). The factor $(2J+1)$, giving the number of sub-states $m_J = -J, -J+1, ..., J-1, J$, increases with J while the exponential factor decreases decreases with J. The product of the two describes the strength of the spectral lines in Fig. 4.13 and is a function of the (excitation) temperature. Fitting the line strengths thus allows to constrain the excitation temperature, by means of the rotational diagram method discussed in Sect. 4.4.

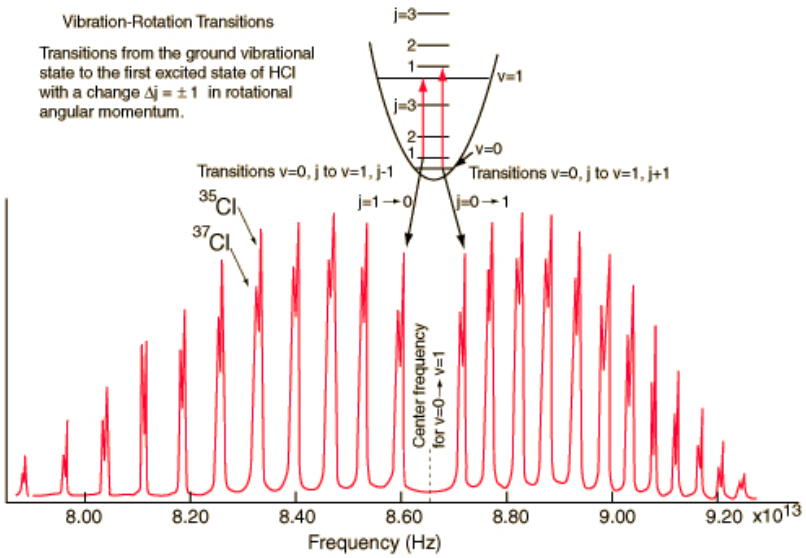


Figure 4.13: *Experimental emission spectrum of the HCl molecule for* $v = 0 \to 1$ *and* $\Delta J = +1$ *and* -1 *transitions. The doubling of the ro-vibrational lines occurs because two chlorine isotopes (*^{35}*Cl and* ^{37}*Cl) were present in the sample to obtain this spectrum. From: hyperphysics.phy-astr.gsu.edu.*

Ro-vibrational interaction

The total vibrational and rotational energy equation (4.36) may be improved using the results Eqs. (4.34) and (4.16). This yields

$$E(v, J) = h\nu_{\text{osc}} \left(v + \frac{1}{2}\right) - \tilde{x} h\nu_{\text{osc}} \left(v + \frac{1}{2}\right)^2 + hcBJ(J+1) - hcDJ^2(J+1)^2. \quad (4.40)$$

Vibrational motion is very fast compared to rotational motion, and as a consequence the rotational states are affected by the vibrational mode the molecule is in. This is handled by introducting two more constants, correcting the B and D values for the vibrational state

$$B_v = B - \alpha \left(v + \frac{1}{2}\right) \qquad \text{and} \qquad D_v = D + \beta \left(v + \frac{1}{2}\right), \quad (4.41)$$

where α and β are the rotation-vibration interaction constants. The final energy expression is obtained by replacing B and D by B_v and D_v in Eq. 4.40. Note that B, α, D, and β are a function of the electronic state E_{elec} of the molecule; these dependencies have not been made explicit in the formulae presented in this chapter. For the ground electronic state of CO, one has $B = 1.922529\,\text{cm}^{-1}$, $D = 6.1193 \times 10^{-6}\,\text{cm}^{-1}$, $\alpha = 0.017507$, and $\beta = 1.0 \times 10^{-9}$.

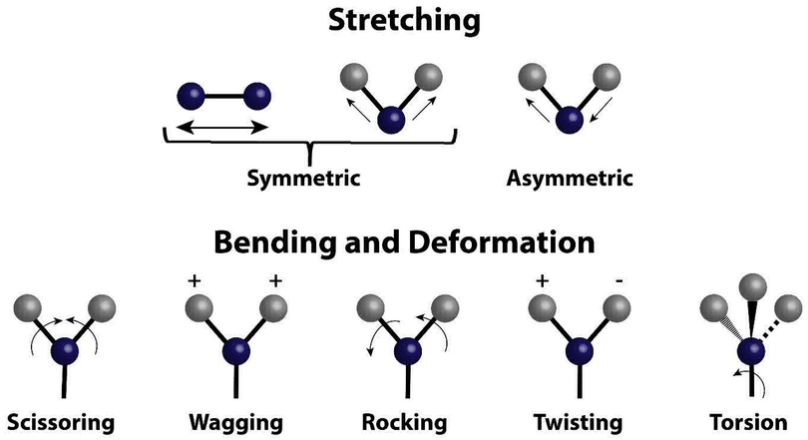


Figure 4.14: *Illustrations of the six vibrational modes in a CH_2 group and the torsion mode of the group as a whole. The H atoms are shown as grey spheres; the C atoms as blue spheres. The carbon atom links the group to the remainder of the molecule. The top left sketch is for a diatomic molecule. A + or - sign indicates motion out of the plane of the paper, towards us or away from us. Bending modes that occur in the plane of the paper (scissoring and rocking) are referred to as in-plane bending modes; those that bring H atoms out of the plane of the paper (twisting and wagging) are out-of-plane bending modes. Stretching modes are symmetric when the two H atoms are in phase; they are asymmetric when they are in anti-phase. For an animation of these modes, see* `http://en.wikipedia.org/wiki/Molecular_vibration`

4.6 Vibrational modes of more complex molecules

A description of the vibrational modes of more complex molecules is beyond the scope of these lectures. However, vibrational modes are very characteristic for the motions of the atoms in the molecular group directly involved but much less sensitive to the structure of the rest of the molecule. This has the benefit that the vibrational frequencies can be related to the fundamental frequency.

Though diatomic molecules only show a stretching motion, complex molecules or distinct groups in complex molecules can vibrate in multiple ways. This is shown in Fig. 4.14 for a methylene group (-CH_2-) that is part of a larger molecule.

In general, a molecule with N atoms has $3N - 6$ normal modes of vibration, but a linear molecule has $3N-5$ such modes. A diatomic molecule thus has one normal mode of vibration. Not all the modes will lead to distinct absorptions. Some of the modes will occur at the same frequency and hence these modes will be degenerate. Others – such as vibrations in homonuclear molecules – may not lead to dipole radiation because the dipole moment does not change during the vibration. So, as an example methane (CH_4) has $3 \times 5 - 6 = 9$ fundamental

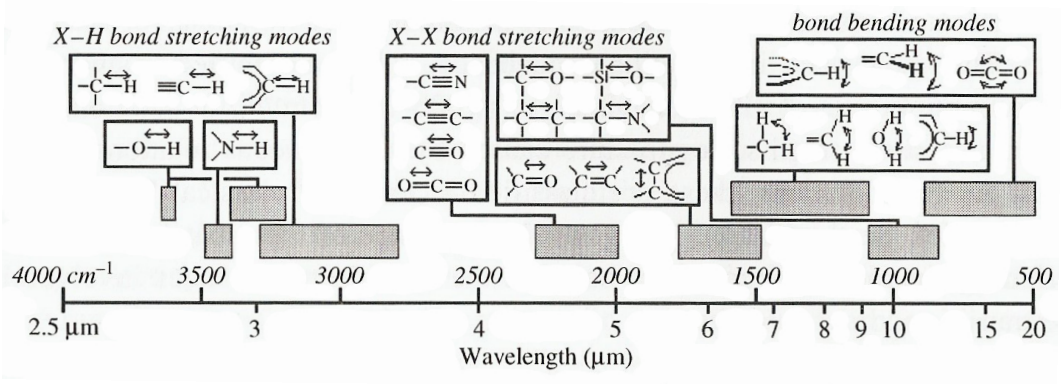


Figure 4.15: *Vibrational frequencies of various molecular groups. The filled boxes indicate the range over which specific molecular groups absorb. The bond type and vibration type of these groups are indicated in the linked boxes. From: D. Hudgins. See Tielens (2005, page 35) for an informative table of vibrational frequencies of molecular groups.*

modes but only the symmetric and asymmetric stretching and bending vibrations of the C−H bonds about the central C atom show up in the IR absorption spectrum: the v_4 mode at 7.66 μm, the v_2 mode at 6.51 μm, and the v_3 mode at 3.31 μm.

Characteristic band positions of various molecular groups are shown in Fig. 4.15. Modes involving motions of H atoms occur at considerably higher frequencies than similar modes involving heavier atoms (a direct result of Eq. 4.28). Thus, H-stretching vibrations occur in the 3 μm region, while stretching motions among (single-bonded) C, N, and O atoms are located around 10 μm. Likewise, when the bond strength increases, from single to double to triple bonds, the vibrations progressively shift to shorter wavelengths.

Exercise 4.1

Consider CO as a linear, rigid rotor.

a) The frequency of the $J = 1 \rightarrow 0$ transition of the main isotopologue of CO ($^{12}C^{15}O$) is 115.3 GHz (see e.g. Table 4.2). What is the internuclear distance in this molecule (in Å)?

b) Assume that the internuclear distance in all CO isotopologues is the same. How large a frequency shift can be expected for the $J = 1 \rightarrow 0$ transition in the $^{13}C^{16}O$ isotopologue relative to 115.3 GHz?

Exercise 4.2

Both H_2 and HD have similar internuclear separation, $r_\circ = 0.741\,\text{Å}$. Assume that the molecules can be approximated as rigid rotators.

a) Calculate $[E(v = 0, J) - E(v = 0, J = 0)]/k$ for H_2 for $J = 1$, $J = 2$, and $J = 3$.

b) Calculate $[E(v = 0, J) - E(v = 0, J = 0)]]/k$ for HD for $J = 1$, $J = 2$, and $J = 3$.

c) Because H_2 has no electric dipole moment, $\Delta J = \pm 1$ transitions are forbidden, and instead the only radiative transitions are electric quadrupole transitions with $\Delta J = 0, \pm 2$. Calculate the wavelength of the $J = 2 \rightarrow 0$ and $J = 3 \rightarrow 1$ transitions of H_2.

d) Because HD has a (small) electric dipole moment, it has (weak) electric dipole transitions. What is the longest-wavelength spontaneous decay for HD in the $v = 0$ vibrational level?

Exercise 4.3

This could be a nice exam question. In order to determine the strength of a rotational transition, we need to know the fractional population of a molecule in rotational level J. In LTE, this is given by (see Eq. 2.70)

$$\frac{n_J}{N} = \frac{(2J+1)}{U(T)} \exp\left(-E_J/kT\right), \tag{4.42}$$

where

$$U(T) = \sum_{J=0}^{\infty} (2J+1) \exp\left(-E_J/kT\right) \tag{4.43}$$

is the partition function (see Eq. 2.71) and N the number density of the molecule. $E(J)$ is given by Eq. (4.12).

Transition	E_u/k [K]	λ [μm]	A_{ul} [s^{-1}]	I_{obs} [erg s^{-1} cm^{-2} sr^{-1}]
0-0 S(1)	1015	17.0348	$4.8\ 10^{-10}$	$6.52\ 10^{-5}$
0-0 S(3)	2504	9.6649	$9.84\ 10^{-9}$	$3.00\ 10^{-4}$
0-0 S(4)	3474	8.0251	$2.64\ 10^{-8}$	$1.24\ 10^{-4}$
0-0 S(5)	4586	6.9095	$5.88\ 10^{-8}$	$1.99\ 10^{-4}$
1-0 S(0)	6471	2.2235	$2.53\ 10^{-7}$	$9.02\ 10^{-5}$
1-0 S(1)	6956	2.1218	$3.47\ 10^{-7}$	$2.36\ 10^{-4}$
1-0 S(2)	7584	2.0338	$3.98\ 10^{-7}$	$6.45\ 10^{-5}$
2-0 S(0)	12095	1.2383	$1.27\ 10^{-7}$	$5.03\ 10^{-5}$
2-0 S(1)	12550	1.1622	$1.90\ 10^{-7}$	$3.05\ 10^{-5}$
2-0 S(2)	13150	1.1382	$2.38\ 10^{-7}$	$3.14\ 10^{-5}$

Table 4.3: *Observed rotational H_2 line strengths towards the young stellar object S106.*

a) Derive an alternative, approximate form of this partition function by replacing the summation with an integral and by taking J to be a continuous variable. This formula should depend only on the temperature T and the rotation constant B.

b) Derive a formula for the value $J = J_{\max}$ at which the maximum population occurs for any specified temperature (so, recover Eq. 4.20). You may do so by assuming the excitation fraction Eq. (4.42) is a continuous function in J, i.e. by taking a similar approach as in a).

c) Find $J_{\max}$ and the corresponding excitation fraction for CO gas at temperatures $T = 10\,\mathrm{K}$ and $T = 100\,\mathrm{K}$. The rotation constant for CO is $B = 1.922529\,\mathrm{cm}^{-1}$.

Exercise 4.4

Vibrational energy separations are larger than rotational ones and are therefore only sensitive to higher temperatures. We consider the molecule CO for which the fundamental frequency of vibration is $2143.3\,\mathrm{cm}^{-1}$. For all diatomic molecules, the vibrational levels v are not degenerate, and so $g_v = 1$.

a) Give a formula that provides the LTE population density ratio of the vibrational state $n(v)$ relative to the vibrational ground state $n(v = 0)$.

b) What proportion of the molecules are in the $v = 1$ vibrational state compared to $v = 0$ at (a) the ISM temperature of $T = 20\,\mathrm{K}$; (b) room temperature, $T = 300\,\mathrm{K}$, and (c) the temperature of a typical M-dwarf star of 3000 K?

Exercise 4.5

In this problem we will use observations of H_2 rotational and ro-vibrational emission lines to construct a rotation diagram and derive the temperature and mass of H_2 in the immediate surroundings of the hot, massive young stellar object S106. This young star is still embedded in its parental molecular cloud and suffers from many tens of magnitudes of optical extinction. We may write for the number density of level u of molecular hydrogen

$$n_u = \frac{g_u}{U(T)} e^{-E_u/kT}\, n(H_2) = \frac{(2J+1)(2I_\mathrm{n}+1)}{U(T)} e^{-hcBJ(J+1)/kT}\, n(H_2), \quad (4.44)$$

where $U(T)$ is the partition function Eq. (2.71) and I_n is the nuclear angular momentum with its statistical weight $g_\mathrm{I} = 2I_\mathrm{n} + 1$. Recall that for para-$H_2$ (even J levels) $I_\mathrm{n} = 0$, while for ortho-H_2 (odd J levels) $I_\mathrm{n} = 1$.

In a previous exercise, we have derived $U(T) = kT/(hcB)$. This result relies on the assumption that the rotational levels are very close to one another. More elaborate analyses show that it is better to approximate $U(T)$ by twice this value, i.e. $U(T) = 2kT/(hcB)$. We will henceforth adopt this better result.

We assume that the spectral lines for the observed transitions of H_2 are optically thin, i.e. that the intensity is given by eq. (4.22) for $J \rightarrow J-2$ transitions. Table 4.3 lists the observed wavelengths and integrated line fluxes for a set of rotational lines observed by the ISO-SWS instrument towards S106. The transitions are indicated by the upper and lower vibrational quantum number and by the rotational quantum number J of the lower level, for the S-branch (i.e. $\Delta J = +2$).

a) What are the rotational quantum numbers J of the upper levels for each of the transitions listed? What are the statistical weights (rotational, nuclear, and total) for these levels?

b) Use the observations of S106 to calculate N_J for each upper level J for the pure rotational lines.

c) Use the above equations to extract a relation between $\log N_\mathrm{J}/g$ and E_J/k.

d) Use the observations of S106 to make a plot of $\log N_\mathrm{J}/g$ as a function of E_J/k. Does the shape of this observed rotation diagram agree with the shape you expect from theory?

e) What is the excitation temperature of the H_2 pure rotational lines (i.e. $v = 0$)? For the $v = 0$ vibrational level, $B = 59.3301\,\mathrm{cm}^{-1}$.

f) What is the total column density of the H_2 gas?

g) Use the data for the ro-vibrational lines (i.e. those with transitions between different vibrational levels) to derive the vibrational excitation temperature. Compare the vibrational excitation temperature to the pure rotation temperature derived above. Explain the difference (if at all) between both temperatures.

5

First steps into the ISM

In this chapter we first make back of the envelope calculations to learn about aspects of, and global properties of, the interstellar medium. The results may reflect order of magnitude estimates only. We then discuss the conditions that should prevail for certain spectral lines to be in LTE in the ISM. We end with a discussion of aspects of the interaction of key components of the ISM, again employing rough estimates. The approximate results obtained in this chapter provide valuable tools for thinking about the physics of the medium in-between the stars.

5.1 Global properties of the interstellar medium

How empty is interstellar space?

How large is the volume occupied by stars with respect to the total volume of our galaxy? For simplicity we assume that all stars are of solar type. The ratio of volumes becomes

$$\frac{V_*}{V_{\rm gal}} \simeq \frac{(4\pi/3)R_*^3\, N_*}{\pi R_{\rm gal}^2\, D_{\rm gal}} \simeq 10^{-22}, \tag{5.1}$$

in which $N_* \simeq 10^{11}$ is the number of stars in the Milky Way; $R_* \simeq 10^{11}$ cm the radius of a star, $R_{\rm gal} \simeq 15$ kpc the radius of the galaxy, and $D_{\rm gal} \simeq 700$ pc the thickness of the stellar disk of the Milky Way. We find a ratio of $\sim 10^{-22}$. Let us compare this to the ratio of the volume of a nucleon (roughly 1 fm or 10^{-13} cm in diameter) relative to the volume of an atom (roughly 1 Å or 10^{-8} cm in diameter). This yields a ratio of $\sim 10^{-15}$, so interstellar space is pretty empty: 10^7 times more so than an atom.

How much mass is in interstellar space?

The ratio of interstellar gas mass to mass in stars in a column perpendicular to the plane of the galaxy, near the Sun, is

$$\frac{H_{\rm gas}\, n_{\rm H}\, m_{\rm H}}{H_*\, n_*\, M_*} \simeq \frac{200\,{\rm pc} \times 1\,{\rm cm}^{-3} \times 1.7\,10^{-24}\,{\rm g}}{700\,{\rm pc} \times 0.1\,{\rm pc}^{-3} \times 2\,10^{33}\,{\rm g}} \simeq 0.08, \tag{5.2}$$

in which $H_{\rm gas}$ and H_* are the typical thickness or scale-heigth of the gas layer and stellar layer (note that these scale-heights in the galaxy are different). $n_{\rm H}$ and n_* are the densities and $m_{\rm H}$ and M_* the masses of respectively the hydrogen atom and of stars. So we find that only ~ 10 percent of the (visible) matter in the solar neighbourhood is interstellar matter.

How much mass is in interstellar gas and how much in interstellar dust?

The ISM consists of gas and of small solid particles, referred to as dust. The *mass ratio of dust to gas* is

$$\frac{(4\pi/3)\, a^3 \rho_{\rm s}\, n_{\rm d}}{m_{\rm H}\, n_{\rm H}} \simeq 0.01 \tag{5.3}$$

in which $n_{\rm d}$ and $n_{\rm H}$ are the volume densities of the dust particles and of the gas atoms respectively. The ratio between these two is $n_{\rm d}/n_{\rm H} \simeq 10^{-12}$. The average radius a of a dust particle in interstellar space is $\simeq 10^{-5}$ cm or 1000 Å or 0.1 μm. The material from which dust grains are composed has a (volumetric mass) density $\rho_{\rm s} \simeq 1\,{\rm g\,cm}^{-3}$. Later, we will discuss that observations have shown that interstellar dust consists of silicates and carbon particles in the form of graphite. In dense regions, water ice (and other volatile dust species) can cover the surfaces of these grains.

5.2 Is the ISM in LTE?

In thermodynamic equilibrium a gas which is kept in a black box has $T_{\rm box} = T_{\rm gas} = T_{\rm radiation}$. This situation is in general not true for matter in the ISM. The interstellar radiation field originates from stars in the galaxy and has a high characteristic temperature (dominated by the luminous OB stars) but is strongly diluted (by the factor given in Eq. 2.8), because the average distance to stars is large. Gas in the ISM is in general colder than the temperature that can be associated with the radiation field. The dust particles have temperatures that are lower than that of the gas. So we have $T_{\rm dust} < T_{\rm gas} < T_{\rm radiation}$.

Let us examine some aspects of LTE in the ISM.

Does the Maxwell velocity distribution hold in the ISM?

An important question is whether or not the velocity distribution of gas particles is Maxwellian under the rather extreme conditions that exist in the ISM. So, can we assign a kinetic temperature to interstellar gas?

Assume that $f(\mathbf{w})$ is the instantaneous velocity distribution. Elastic collisions between atoms and ions in the gas lead to equipartition of energy and so to the Maxwell velocity distribution $f^\circ(\mathbf{v})$, where the superscript $\circ$ indicates that f is Maxwellian. The rate of change of the velocity distribution toward a Maxwellian distribution is given by

$$\left(\frac{\partial f}{\partial t}\right)_{\rm elas} = \frac{f^\circ(\mathbf{v}) - f(\mathbf{w})}{t_{\rm elas}}, \tag{5.4}$$

where $t_{\rm elas}$ is the typical timescale of elastic collisions. Interactions between gas particles in which the atoms and/or ions get excited as a result of collisions lead to a reduction of the number of gas particles with a kinetic energy $\frac{1}{2}mw^2 > E_{\rm exc}$, where $E_{\rm exc}$ is the excitation energy. Therefore, the rate of change toward establishing the velocity distribution that reflects the missing high kinetic energy particles in the collision partner population is

$$\left(\frac{\partial f}{\partial t}\right)_{\rm exc} = -\frac{f(\mathbf{w})}{t_{\rm exc}}, \tag{5.5}$$

where $t_{\rm exc}$ is the typical timescale for collisions that lead to an excitation. In a steady state situation we must have $\partial f/\partial t = 0$, therefore

$$\frac{f^\circ(\mathbf{v}) - f(\mathbf{w})}{f(\mathbf{w})} = \frac{t_{\rm elas}}{t_{\rm exc}}. \tag{5.6}$$

Let us consider a medium in which hydrogen is neutral, so conditions typical for the CNM (see Table 1.1). The most important elastic collision process is that of H I particles colliding with each other. A rough estimate for the typical timescale for such collisions is $t_{\rm elas} \simeq (n_{\rm H}\,\sigma_{\rm elas}\,v_{\rm H})^{-1}$. The dominant inelastic scattering process is that of collisional excitation of the ground electronic state $1s^2 2s^2 2p^1\ {}^2P^\circ$ of C II, which contains two fine-structure levels (see Fig. 5.1), ${}^2P^\circ_{1/2}$ and ${}^2P^\circ_{3/2}$. The ${}^2P^\circ_{3/2}$ fine-structure level can be excited by collisions of ${}^2P^\circ_{1/2}$ with H I. A rough estimate for the typical timescale for such collisions is $t_{\rm exc} \simeq (n_{\rm CII}\,\sigma_{\rm exc}\,v_{\rm H})^{-1}$.

In H I gas we can approximately equate $\sigma_{\rm elas} \simeq \sigma_{\rm exc}$ and $n({\rm C\,II}) \simeq X_{\rm C}\,n_{\rm H}$, where $X_{\rm C}$ is the number fraction of carbon relative to hydrogen, since the carbon which is in the gas phase will be almost entirely singly ionized. Then we find

$$\frac{f^\circ(\mathbf{v}) - f(\mathbf{w})}{f(\mathbf{w})} \simeq X_{\rm C} \simeq 3\,10^{-4}. \tag{5.7}$$

So, fortunately the deviations from a Maxwellian velocity distribution will be very small, simply because the number of elastic collisions dominates.

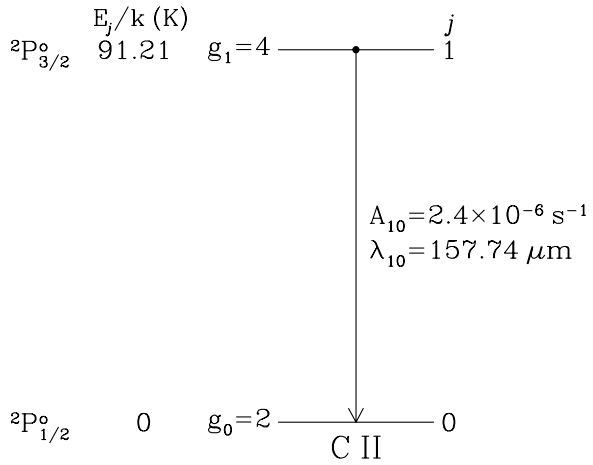


Figure 5.1: *Fine structure levels of C* II. *From: Bruce Draine*, Physics of the interstellar and intergalactic medium.

In H II regions the gas will approach the Maxwell velocity distribution even better because $\sigma_{\rm elas} \gg \sigma_{\rm exc}$. This is due to the long-range character of the coulomb force between charged particles.

Is the 21 cm transition of hydrogen in LTE?

Consider the 21 cm transition of neutral hydrogen atoms in interstellar space. We adopt $n_{\rm H} \simeq 1$ cm^{-3} and $T \simeq 70$ K. Collisions among H I particles are capable of populating the two hyper-fine-structure levels according to the Boltzmann equation if the collisional de-excitations are much more frequent than radiative de-excitations, i.e. if $t_{\rm coll} \ll t_{\rm rad}$. If we approximate the collisional cross-section for hydrogen atoms by $\sigma_{\rm H} = \pi a_0^2$, where $a_0 = 0.53 \times 10^{-8}$ cm is the Bohr radius, we find:

$$t_{\rm coll} \simeq \frac{1}{n_{\rm H}\, \sigma_{\rm H} < v_{\rm H} >} \simeq \frac{2100}{n_{\rm H}} \quad \text{yrs}, \tag{5.8}$$

where we have used the average relative speed of hydrogen atoms of $1.7\,{\rm km\,s^{-1}}$ for $T = 70$ K, as computed below Eq (2.66).

The radiation timescale follows from the transition probability $A_{21} = 2.87 \times 10^{-15}\,{\rm s^{-1}}$ for spontaneous de-excitation in the 21 cm line. As this value is small,

$$t_{\rm rad} \simeq \frac{1}{A_{21}} \simeq 1.1 \times 10^7 \quad \text{yrs}. \tag{5.9}$$

is large. So, the level populations of the hyper-fine levels in the electronic ground state of hydrogen are distributed according to Boltzmann ($n_2/n_1 \simeq g_2/g_1 = 3$ as for typical conditions $h\nu/kT \ll 1$) for gas densities $n_{\rm H} \gtrsim 10^{-4}$ cm^{-3}. This is virtually everywhere in interstellar space.

Is the first rotational excited state of para-H_2 in LTE?

The first rotationally excited state of para-H_2 connects to the ground state of molecular hydrogen. The energy levels of rotational states are given by Eq. (4.12), which we repeat here for convenience

$$E_J = hc\,BJ(J+1), \tag{5.10}$$

in which $B = 60.80\,\mathrm{cm}^{-1}$ is the rotational constant. The energy of the transition is at

$$\Delta E = E(J=2) - E(J=0) = 6hcB = 7.25 \times 10^{-14}\ \mathrm{erg} = k\,[540\,\mathrm{K}], \tag{5.11}$$

Typical temperatures observed in regions of molecular gas in the ISM are 10–40 K. Collisions can therefore not (significantly) excite H_2. The $J = 2$ state of para-H_2 is thus not in LTE.

This result has a bigger implication. As the lowest exited J state of hydrogen is only very poorly populated, it does not in general emit much radiation. Moreover, as H_2 is a homonuclear molecule, therefore it has no permanent electric dipole moment and hence its dipole transitions are forbidden (see Sect. 4.2). As already pointed out, its molecular structure is essentially a symmetric 'dumb-bell'. The combined low excitation of the $J = 2$ level and the forbidden nature of the transition cause the H_2 *emission* line to be very weak.

Though H_2 can be detected via its UV *absorption* lines (the same lines that are involved in its destruction), this is only on the lines of sight to a few suitably positioned bright, hot, background stars. These lines of sight are too few and far between to make this an effective means of conducting either a global search for, or a survey of, H_2. Since H_2 is so hard to observe, carbon monoxide (CO) is used as a tracer of molecular gas.

Is the CO molecule in LTE in molecular clouds?

The first rotationally excited level of CO has energy

$$\Delta E = E(J=1) - E(J=0) = 2hcB = 7.64 \times 10^{-16}\ \mathrm{erg} = k\,[5.5\,\mathrm{K}], \tag{5.12}$$

where $B = 1.922529\,\mathrm{cm}^{-1}$. From $\lambda = hc/\Delta E$ we find that the transition is at 2.6 mm. Recall that for permitted rotational transitions $\Delta J = 1$. This line can be easily excited at the low temperatures observed in regions of molecular gas in the ISM. For a temperature $T = 30\,\mathrm{K}$ and densities $n_\mathrm{c} > 10^4\,\mathrm{cm}^{-3}$ we find that $t_\mathrm{coll} < t_\mathrm{rad}$. For these conditions the line is in LTE.

Are also higher rotational states in LTE? The Einstein A-coefficients for the rotational transitions is given by

$$A_{J,J-1} = \frac{2^9\pi^4 B^3}{3h}\,d^2\,\frac{J^4}{2J+1}, \tag{5.13}$$

in which $d = 0.11 \times 10^{-18}\,\mathrm{Fr{\cdot}cm}$ is the dipole moment of the molecule. This results in transition probabilities of the rotational transitions of the CO molecule given by

$$A_{J,J-1} \simeq 2.3\,10^{-7}\,\frac{J^4}{2J+1}\quad \mathrm{s}^{-1}. \tag{5.14}$$

The life time of an excited CO rotational level is the inverse of the Einstein A coefficient

$$t_J \propto A_{J,J-1}^{-1} J^{-3}. \tag{5.15}$$

So we find that for the high J rotational levels the depopulation of energy levels due to radiation is increasingly more rapid, and at some value of J this radiative depopulation can no longer be compensated by collisional or radiative population. In a typical molecular cloud we find a Boltzmann distribution for the lower J levels according to the kinetic temperature of the gas, but in the higher energy levels the level populations deviate significantly from LTE.

5.3 Non-LTE in the interstellar medium

We may conclude from the examples in the previous section that for transitions among levels that are close to the ground level and for which the energy difference between the lower and upper level is small, collisional transitions may dominate. Examples of such cases may be fine structure lines. However, when the energy difference between lower and upper level is large, so typically for transitions in the optical or ultraviolet part of the spectrum, collisions between gas particles will not always be energetic enough to cause an excitation of the electron. This will inevitably lead to a situation of NLTE, i.e. a situation in which we can no longer rely on the Boltzmann and Saha equations, Eqs. (2.67) and (2.72). If that is so, we need to solve the *statistical equilibrium equations* to obtain the state of the gas

$$\sum_{u \neq l} (n_u P_{ul} - n_l P_{lu}) = 0, \tag{5.16}$$

where l refers to the lower state and u to the upper state. P_{lu} and P_{ul} are the rates at which electrons transit from the lower to the upper state and vice versa, each of which is the sum of a radiative rate and a collisional rate. To keep things simple, we will here consider the statistical equilibrium equation for a system consisting of two bound levels. In Section 8.2, we will formulate the rates that describe transitions between a bound and a free level.

Two level system

We consider a schematic atomic or molecular model consisting of two levels between which radiative and collisional transitions can occur. This model is obviously very incomplete, but it nevertheless provides a fairly good description of the real situation for some lines. Two processes can result in a transition of the electron from a lower state l to an upper state u. The number of excitations per second per cm^3 caused by these processes are

$$\begin{aligned} &- \text{Radiative excitations} \quad : \quad \frac{\mathrm{d}n_l}{\mathrm{d}t} = n_l R_{lu} = n_l B_{lu} \bar{J} \\ &- \text{Collisional excitations} \quad : \quad \frac{\mathrm{d}n_l}{\mathrm{d}t} = n_l C_{lu} = n_l n_\mathrm{c} q_{lu} \end{aligned}$$

Three processes can result in the transition of the electron from the upper state to the lower state. The number of such de-excitations (also in cm^{-3} s^{-1}) are

$$\begin{aligned}
&- \text{Spontaneous emissions} \quad : \quad \frac{\mathrm{d}n_u}{\mathrm{d}t} = n_u\, R_{ul}^{\mathrm{spon}} = n_u\, A_{ul} \\
&- \text{Stimulated emissions} \quad : \quad \frac{\mathrm{d}n_u}{\mathrm{d}t} = n_u\, R_{ul}^{\mathrm{stim}} = n_u\, B_{ul}\, \bar{J} \\
&- \text{Collisional de-excitations} \quad : \quad \frac{\mathrm{d}n_u}{\mathrm{d}t} = n_u\, C_{ul} \quad = n_u\, n_{\mathrm{c}}\, q_{ul}
\end{aligned}$$

In the above expressions A_{ul}, B_{lu} and B_{ul} are the Einstein coefficients, and R and C (with indices lu and ul) the radiative and collisional rate. The dimension of A_{ul} is sec^{-1}, that of B_{lu} and B_{ul} is erg^{-1} cm^2 hz; the dimension of R and C is sec^{-1}. $\bar{J}$ (which could have been given the indices lu and ul for completeness) is the profile function averaged mean intensity

$$\bar{J} = \bar{J}_{lu} = \bar{J}_{ul} = \int_0^\infty \phi(\nu)\, J_\nu\, d\nu, \tag{5.17}$$

where $\phi(\nu)$ is the profile function. The profile function is normalized, i.e.

$$\int_0^\infty \phi(\nu)\, d\nu \equiv 1, \tag{5.18}$$

and sharply peaked, therefore $\bar{J} \simeq J_\nu$, i.e. it is almost equal to the mean intensity at frequency ν at line center. In stating that $\bar{J}_{lu} = \bar{J}_{ul}$ we assume that the profile function for radiative excitations is the same as for stimulated emissions. The situation that the profile functions of all three radiative processes are the same (so also that of spontaneous emissions) is known as *complete redistribution* and is realized in TE.

The quantities q_{lu} en q_{ul} give the product of the (velocity dependent) cross-section for collisions and the velocity of the colliding partner, integrated over the relevant part of the Maxwell velocity distribution, in cm^3 s^{-1}. Let us call this function the *collision strength*. We take as a collision partner particles of type 'c' (collision) that have a number density n_{c} cm^{-3}. In ionized H II regions, free electrons are typically the collision partner. In molecular clouds, the H_2 molecule will often be the collision partner. In H I regions it will be the hydrogen atom.

Now that we have formally introduced the Einstein coefficients, we also formally mention that the *lifetime* (in seconds) of an excited level is related to the Einstein A_{ul} coefficient through

$$\tau_u = \frac{1}{\Sigma_l A_{ul}}. \tag{5.19}$$

Some numbers: for permitted transitions of hydrogen the values of $A_{ul} \approx 10^4 - 10^8$ sec^{-1}. From this we get that $\tau_u \sim 10^{-4} - 10^{-8}$ sec, so these excited levels have a very short lifetime.

Statistical equilibrium between upper and lower level implies that the time average of the population numbers of the levels do not change. This requires that

$$n_l \left(B_{lu}\, \bar{J} + n_{\mathrm{c}}\, q_{lu}\right) = n_u \left(A_{ul} + B_{ul}\, \bar{J} + n_{\mathrm{c}}\, q_{ul}\right) \tag{5.20}$$

We can rewrite this equation by making use of the relations between the Einstein coefficients

$$g_l B_{lu} = g_u B_{ul} \tag{5.21}$$

$$A_{ul} = \frac{2h\nu^3}{c^2} B_{ul} \tag{5.22}$$

To make efficient use of the Einstein relations, i.e. to conveniently express B_{lu} and B_{ul} in A_{ul}, we introduce the *photon occupation number*

$$n_\gamma \equiv J_\nu \Big/ \frac{2h\nu^3}{c^2} \,, \tag{5.23}$$

such that when $\bar{J} \simeq J_\nu = B_\nu \rightarrow n_\gamma = \left(\mathrm{e}^{h\nu/kT} - 1\right)^{-1}$. For the radiative excitation processes we can now write:

$$n_l \, B_{lu} \, \bar{J} = n_l \, \frac{g_u}{g_l} B_{ul} \, J_\nu = n_l \, \frac{g_u}{g_l} \, A_{ul} \, \frac{c^2}{2h\nu^3} \, J_\nu = n_l \, \frac{g_u}{g_l} \, A_{ul} \, n_\gamma, \tag{5.24}$$

and for the stimulated emission processes

$$n_u \, B_{ul} \, \bar{J} = n_u \, B_{ul} J_\nu = n_u \, A_{ul} \, \frac{c^2}{2h\nu^3} \, J_\nu = n_u \, A_{ul} \, n_\gamma. \tag{5.25}$$

Using these results, statistical equilibrium equation (5.20) can be written as

$$\frac{n_u}{n_l} = \frac{n_\mathrm{c} \, q_{lu} + (g_u/g_l) \, n_\gamma \, A_{ul}}{n_\mathrm{c} \, q_{ul} + (n_\gamma + 1) A_{ul}} \tag{5.26}$$

We may consider a few simple limiting cases that are relevant for the conditions in the ISM. We first ignore the radiation field, and second we will ignore collisions. In the first case we expect to get solutions close to LTE, while in the second case strong NLTE effects may occur.

Ignoring the radiation field

Let's first consider the case where radiation is ignored, i.e. $n_\gamma = 0$. The statistical equilibrium equation reduces to

$$\frac{n_u}{n_l} = \frac{n_\mathrm{c} \, q_{lu}}{A_{ul} + n_\mathrm{c} \, q_{ul}} = \frac{n_\mathrm{c} \, q_{lu}}{A_{ul}} \left(\frac{1}{1 + n_\mathrm{c} \, q_{ul}/A_{ul}} \right) \tag{5.27}$$

Again, two limiting cases can be considered, with high and low density of the collision partners. In the low density limit we find

$$n_\mathrm{c} \, q_{ul} \ll A_{ul} \quad \rightarrow \quad \frac{n_u}{n_l} = \frac{n_c \, q_{lu}}{A_{ul}}. \tag{5.28}$$

The balance between n_l and n_u is set by collisional excitations (even though the density of the medium is low) and spontaneous recombinations (that dominate over collisional de-excitations).

Similarly, for the high density limit we get

$$n_{\rm c}\, q_{ul} \gg A_{ul} \quad \rightarrow \quad \frac{n_u}{n_l} = \frac{q_{lu}}{q_{ul}} = \frac{g_u}{g_l}\, \exp\left[-h\nu/kT\right]. \tag{5.29}$$

The last equality links q_{lu}/q_{ul} to the Boltzmann distribution. In a medium where the state of the gas (in our example the densities of n_l and n_u) is determined by collisions only, and such collisions are frequent, it is the velocity distribution $f(\mathbf{v})$ and density $n_{\rm c}$ of the collision partner particles that determines the outcome. In the ISM, the velocity distribution is according to Maxwell, therefore the distribution only depends on the (local) temperature and hence will be in (local) TE. In short

$$n_l^*\, C_{lu} = n_u^*\, C_{ul} \qquad \text{or} \qquad n_l^*\, q_{lu} = n_u^*\, q_{ul}. \tag{5.30}$$

It thus follows that the population ratio n_u/n_l must be according to Boltzmann.

We can now introduce an important concept in ISM gas analysis, namely that of the critical density

$$n_{\rm cr} \equiv \frac{A_{ul}}{q_{ul}}. \tag{5.31}$$

The physical meaning of the critical density is clear: for $n_{\rm c} < n_{\rm cr}$ collisional de-excitations will be unimportant, and most de-excitations will be spontaneous, resulting in the emission of a photon. The strength of such emission lines will therefore be a good measure of density (i.e. of $n_{\rm c}$ in Eq. 5.28). For densities much above $n_{\rm cr}$ de-excitation will be mainly through collisions. We expect to approach LTE. Line strengths (though not the dominant de-excitation mechanism, spontaneous de-excitations may still occur) will provide a measure of the local kinetic temperature (i.e. of T in Eq. 5.29). The critical density is determined by the quantum-mechanical properties of the transition under consideration: small Einstein A_{ul} values result in low critical densities, i.e. densities in the range typically observed in the ISM. This is why forbidden lines (including fine structure lines) are so important for ISM studies. An overview of important forbidden lines is given in Fig. 5.2. Note that study of these lines allows to probe a wide range of temperatures and densities. Forbidden lines are not important in stellar atmospheres, since the densities are usually far above the critical density for any forbidden line (clearly permitted lines are abundant in stellar spectra). We can generalize the expression for the critical density to multilevel systems where the critical density now compares radiative transitions to all lower levels with the collisional transitions to all levels

$$n_{\rm cr} = \frac{\Sigma_{l<u} A_{ul}}{\Sigma_{l\neq u} q_{ul}}. \tag{5.32}$$

The principle is the same: LTE ensues when the density is larger than the critical density. In answering the questions in section 5.2, we computed the time scale for collisional processes

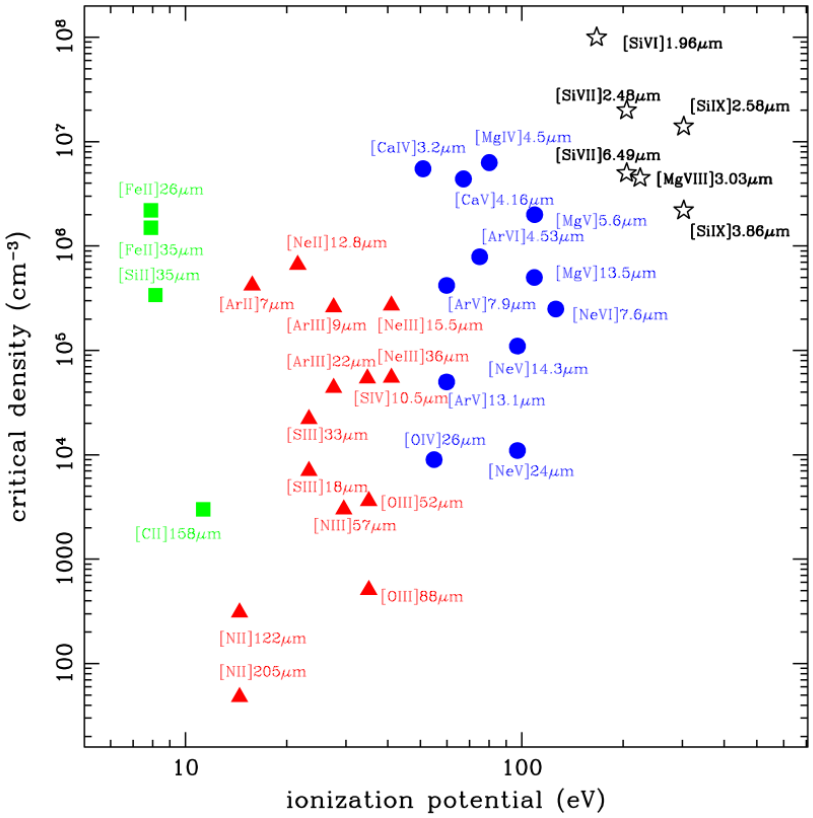


Figure 5.2: *Critical density for collisional de-excitation versus ionization potential of infrared fine-structure lines. Note that a wide range in temperatures and densities is accessible. From: Spinoglio et al. 2012, ApJ 745, 171.*

and compared these to the timescale for spontaneous de-excitation (see Eqs. 5.8 and 5.9). This is an alternative way of assessing whether the medium is in LTE or not. For $t_{\rm coll} < t_{\rm rad}$, LTE will be valid as collisions will (start to) dominate the excitation and de-excitation processes. Note that the critical density is also equivalent to

$$n_{\rm cr} \equiv \frac{A_{ul}}{q_{ul}} = \frac{t_{\rm coll}}{t_{\rm rad}}\, n_{\rm c}, \tag{5.33}$$

where $n_{\rm c}$ is the colliding partner.

Ignoring collisions

Now let us consider the case that collisions are not important, but radiation is. In the equation for statistical equilibrium we set $n_{\rm c} = 0$. We find

$$\frac{n_u}{n_l} = \frac{g_u}{g_l} \frac{n_\gamma}{n_\gamma + 1} \tag{5.34}$$

We can also switch off radiation: $n_\gamma \to 0$. We obtain the trivial result that $n_u/n_l \to 0$, all particles are in the ground state and stay there.

More interesting is to assume the radiation field to be a Planck function, i.e. $J_\nu = B_\nu(T_{\rm rad})$, such that $n_\gamma = \left(\mathrm{e}^{h\nu/kT_{\rm rad}} - 1\right)^{-1}$. We find

$$\frac{n_u}{n_l} = \frac{g_u}{g_l} \exp\left[-h\nu/kT_{\rm rad}\right]$$

The physical meaning of this limiting case is that the radiation field thermalizes the level populations. These level populations are forced to correspond to the temperature of the radiation field, referred to as the *radiation temperature* $\mathrm{T_{rad}}$. The value of $\mathrm{T_{rad}}$ can be quite different from the actual kinetic temperature of the gas. It is clear that the impact of the radiation field on level populations can be significant. This is especially true in cases where the average photon energy of the radiation field is well matched to the energy difference between upper and lower levels of a particular transition, or class of transitions. We mention a few examples of cases where the diffuse and/or local radiation field is important and influences level populations

- Fine structure lines have their transitions in the infrared and will couple well to the radiation field of circumstellar and/or interstellar dust. Recall that stellar photospheres emit most photons at short wavelengths - these can be ignored.
- Light molecules, such as H_2, CH, OH, have strong transitions in the mid-infrared and will also couple well to the dust radiation field.
- Heavy molecules, such as CO and SiO, have rotational transitions in the millimeter wavelength range and so they will be sensitive to the 3 K cosmic background radiation field.

Optical depth effects

So far we have assumed the gas to be optically thin, i.e. we have ignored optical depth effects. There are however many examples of conditions where this is not the case, for instance for CO rotational line emission from molecular clouds. The problem quickly becomes rather

complicated because of the inherent non-local nature of the radiation field. So, the level population equations and radiative transfer equations have to be solved simultaneously. We can make life simpler by introducing the concept of *escape probability by direct flight* β.

We assume that photons that are produced locally can only be absorbed locally. If they are *not* absorbed locally they escape to infinity by direct flight. This can for instance be the case when there are (large) velocity gradients in the gas, which may occur in expanding, rotating, or collapsing clouds. We further assume the local region of the cloud is homogeneous (so the populations are the same everywhere) and that the cloud is not irradiated by an external radiation field. We then have the situation in which the photons that are absorbed are those photons that do not escape, i.e.

$$\bar{J}\,(n_l\,B_{lu} - n_u\,B_{ul}) = (1-\beta)\,n_u\,A_{ul}, \tag{5.35}$$

where in the left-hand-part we have corrected the absorptions for a (negative) contribution by stimulated emissions. This is customary practice as the photons produced in stimulated emission events will have the same frequency and direction as the photons responsible for these emissions.

Alternatively, we could have expressed the profile averaged radiation field as

$$\bar{J} = (1-\beta)\,S_\nu, \tag{5.36}$$

where S_ν is the local line source function for the transition between levels l and u. That is, $\bar{J}$ is equal to the locally emitted radiation corrected for the fraction of the light that escapes without interacting with the local medium. This should be equivalent to Eq. (5.35). The line emission and line extinction coefficient are given by

$$\eta_\nu^\ell = \frac{h\nu_{lu}}{4\pi}\,n_u\,A_{ul}\,\phi_\nu \tag{5.37}$$

$$\chi_\nu^\ell = \frac{h\nu_{lu}}{4\pi}\,[\,n_l\,B_{lu}\,\phi_\nu - n_u\,B_{ul}\,\phi_\nu\,]\,. \tag{5.38}$$

The energy emitted in all directions per cm^3 per second at all frequencies relevant for the transition is $h\nu_{lu}\,n_u A_{ul}$. To get the emitted energy per hz we need to multiply with the profile function ϕ_ν (which has dimension hz^{-1}), and, finally, to get the emission per unit solid angle we need to divide by the total solid angle $\Omega = 4\pi$. The total energy that is absorbed in a volume element dV in a time interval dt by radiative excitation is

$$\begin{aligned} dE_\nu^{\rm tot} &= h\nu_{lu}\,n_l\,B_{lu}\,\bar{J}_{lu}\,dV\,dt \\ &= h\nu_{lu}\,n_l\,B_{lu}\left[\int\left\{\frac{1}{4\pi}\oint I_\nu\,d\omega\right\}\phi_\nu\,d\nu\right]\,dV\,dt \end{aligned} \tag{5.39}$$

where the last equality at the right-hand side follows (5.17). So, the energy dE_ν that is absorbed in a volume $dV = dOds$ in a bandwidth $d\nu$ in a time interval dt, from an incident beam of opening angle $d\omega$ and specific intensity I_ν is

$$dE_\nu = \frac{h\nu_{lu}}{4\pi}\,n_l\,B_{lu}\,\phi_\nu\,I_\nu\,dO\,ds\,d\omega\,d\nu\,dt \tag{5.40}$$

$$= \chi_\nu\,I_\nu\,dO\,ds\,d\omega\,d\nu\,dt \tag{5.41}$$

This recovers the first right-hand-side term in Eq. (5.38). The second right-hand-side term is the correction for stimulated emission that can be derived in a similar manner.

For the line source function we now find

$$S_\nu^\ell = \frac{\eta_\nu^\ell}{\chi_\nu^\ell} = \frac{n_u\, A_{ul}}{n_l\, B_{lu} - n_u\, B_{ul}}, \tag{5.42}$$

where we have adopted that the profile functions for spontaneous emissions, radiative emissions and stimulated emissions are identical and given by ϕ_ν. Substitution of Eq. (5.42) in (5.36) recovers Eq. (5.35), as announced.

Combining Eqs. 5.35 and 5.20 we find

$$n_l\, n_c\, q_{lu} = n_u\, n_c\, q_{ul} + n_u\, \beta\, A_{ul}, \tag{5.43}$$

which is an expression similar in shape as Eq. (5.27). Following the discussion from there, we can again formulate an expression for the critical density

$$n_{\rm cr} = \frac{\beta\, A_{ul}}{q_{ul}}. \tag{5.44}$$

This shows that the inclusion of optical depth effects by means of an escape probability formalism lowers the critical density, because $0 \leq \beta \leq 1$. When the medium is no longer optically thin, photons may get 'trapped'. Equation (5.35) essentially states that with a probability β a spontaneous radiative de-excitation is 'on the spot' followed by a radiative excitation, as if nothing happened. The only 'real' spontaneous de-excitations occur at the lower rate βA_{ul}. It is this rate to which the collisional down channel has to be compared. In other words, LTE is reached at a lower gas density (compared to the optically thin situation described in Eq. 5.31).

5.4 Examples of interactions between the components constituting the ISM

Interaction between dust and radiation - heating of dust grains by radiation

Though in reality the interstellar radiation field consists of several components (see Fig. 1.6 and Chapter 6), let us characterize it as a diluted Planckian radiation field with a radiation temperature of about 10^4 K. The total mean intensity is

$$J = W\, B(T_{\rm rad}), \tag{5.45}$$

where 'total' implies that we have integrated over all frequencies. The energy density of this radiation field is (see Eq. 2.11)

$$u = \frac{4\pi}{c}\, W\, B(T_{\rm rad}) = W \frac{4\pi}{c}\, \frac{\sigma T_{\rm rad}^4}{\pi}. \tag{5.46}$$

Note that the dilution factor is to be interpreted as a mean of the dilution of the radiation of many stars, therefore we do not suggest to adopt the distance to the nearest star. A typical mean dilution factor for the ISM is $W \simeq 10^{-14}$, a number that is related to the average distance in-between the stars in the galaxy that contribute most to the radiation field. Adopting $T_{\rm rad} = 10\,000$ K yields an energy density of radiation $u \simeq 10^{-12}$ erg cm^{-3} $\simeq 0.6$ eV cm^{-3} (1 erg = 6.2415×10^{11} eV; see also Table 1.3). Physically this means that the radiation energy density is very low, but that the energy per photon is high: there are very few photons per unit volume, but those that are present are quite energetic.

If we assume that dust particles radiate as black bodies, we can derive the thermal equilibrium temperature that they should obtain through absorption of photons from the interstellar radiation field. The amount of energy absorbed per unit time is $4\pi\, J$, i.e. the mean intensity integrated over all solid angles, times the cross section of the particle. We find that

$$E_{\rm abs} = 4\pi\, J(T_{\rm rad})\, \pi a^2 = 4\pi\, W\, B(T_{\rm rad})\, \pi a^2 = 4\, W\, \sigma T_{\rm rad}^4\, \pi a^2. \tag{5.47}$$

Note that as the stars are all at large distance, $4\pi\, J = \mathcal{F}$, where $\mathcal{F}$ is the flux that illuminates the particle (to see this, combine Eqs. 2.7 and 2.20).

Similarly, the dust grain emits black body radiation at a dust temperature $T_{\rm d}$. The energy per unit time radiated in this way is

$$E_{\rm em} = 4\pi a^2\, \sigma T_{\rm d}^4 \tag{5.48}$$

Equating the absorbed and emitted energy

$$T_{\rm d} \simeq W^{1/4}\, T_{\rm rad} \tag{5.49}$$

Substitution yields 3 K. In reality the grains in interstellar space are warmer than 3 K, they are found to be in the range $\sim 15{-}25$ K. This is because of other contributions to the interstellar radiation field, but mainly because dust grains do not radiate as black bodies. We will return to the latter subject when we discuss interstellar dust.

Interaction between dust and radiation - extinction of light by dust grains

The density of dust particles in the ISM is $n_{\rm d} \simeq 10^{-12}$ cm^{-3} and the average radius of a grain is $a_{\rm d} \simeq 10^{-5}$ cm. Using these numbers, we find for the mean free path of a photon for dust absorption through the ISM (see Eq. 2.45):

$$\ell = \frac{1}{\pi a_{\rm d}^2\, n_{\rm d}} \simeq 1\,{\rm kpc}. \tag{5.50}$$

This means that we cannot observe objects in the plane of our galaxy at distances much larger than a few kpc. We have assumed in this calculation that the cross-section for dust absorption is the geometric cross section. This is a reasonable assumption for wavelengths smaller than the typical size of the particles. At long, far-infrared and radio wavelengths, the cross-section for radiation to be absorbed by dust is much lower, and we can observe the entire galaxy.

Interaction between dust and gas - gas particles decelerating dust grains

The *stopping distance* of a moving dust particle is defined as the distance a dust grain can travel before it has swept up its own mass by (sticking) collisions between the dust grain and gas particles (mostly molecular hydrogen in molecular clouds). We may write for this stopping distance

$$\ell_{\rm d} = \frac{4\pi a^3\,\rho_{\rm s}/3}{\pi a^2\,n_{\rm H}\,{\rm m_H}} = \frac{4\,a\,\rho_{\rm s}}{3\,n_{\rm H}\,m_{\rm H}} \simeq \frac{5}{n_{\rm H}} \quad {\rm pc}, \tag{5.51}$$

where $a \simeq 10^{-5}$ cm is the grain radius and $\rho_{\rm s} \simeq 2\,{\rm gr\,cm^{-3}}$ is the mass density of the dust particles. In clouds with ${\rm n_H} > 10\ {\rm cm^{-3}}$ we find that $\ell_d < 0.5$ pc. This is smaller than the typical dimension of an interstellar cloud ($\sim$ 1–10 pc). Gas and dust therefore interact efficiently and as a result are well mixed. It is not easy to spatially separate the two components, which would happen if the frictional force between gas and dust would be weak. We conclude that in the ISM gas and dust are collisionally coupled.

The collision timescale for a hydrogen atom and a dust grain is

$$t_{\rm H-dust} = \frac{1}{n_{\rm d}\,\sigma_{\rm d}\,<v_{\rm H}>} \simeq \frac{10^9}{n_{\rm H}} \quad {\rm yr}, \tag{5.52}$$

where we again have used that $n_{\rm d}/n_{\rm H} \simeq 10^{-12}$ and an average velocity of hydrogen particles of 1.7 ${\rm km\,s^{-1}}$(see Eq. 2.66). For the collisional cross section of the grains we adopted $\sigma_{\rm d} = \pi a^2$, where $a = 10^{-5}$ cm is a typical grainsize.

So, on average one hydrogen atom has collided only 10 times with a dust particle during the lifetime of the galaxy assuming $n_{\rm H} = 1\,{\rm cm^{-1}}$. In denser environments, such as molecular clouds, collisions are much more frequent. These collisions are important because this is the only way in which molecular hydrogen H_2 can be made. Collisions between two hydrogen atoms do not form H_2 because the excess binding energy cannot be radiated away fast enough. The reason for this is that H_2 is a so-called homo-nuclear molecule which shows only dipole-forbidden rotational transitions. This problem can be circumvented by introducing a third particle that can carry away the excess energy. At very high gas densities (which do not occur in the ISM), a three-body collision of H atoms can lead to H_2 formation. In ISM conditions dust grains act as a third body, which can absorb the binding energy of the molecule. This shows the importance of *grain surface reactions* for interstellar chemistry.

$$2{\rm H} + {\rm dust} \rightarrow {\rm H_2^\star} + {\rm dust} \rightarrow {\rm H_2} + {\rm dust} + h\nu, \tag{5.53}$$

where the $\star$ sign indicates that the hydrogen molecule is excited.

The formation of H_2 in clouds becomes efficient at densities above about $n_{\rm H} \gtrsim 300\ {\rm cm^{-3}}$. These clouds are called molecular clouds.

Interaction between gas and radiation

We can estimate the energy density associated with the thermal motions of gas in H I regions. Adopting an average gas density of $n_{\rm H} \simeq 1$ cm^{-3} and a temperature of T $\simeq$ 100 K, we find for the thermal energy density $E_{\rm th} = 3\, n_{\rm H}\, kT/2 \simeq 2 \times 10^{-14}$ erg cm^{-3} or 0.01 eV cm^{-3}, which is a lot smaller than the energy density of the interstellar radiation field. This number also implies that collisions among hydrogen particles will not lead to ionization, for which 13.6 eV is needed per event.

The mean free path of a photon with energy $h\nu > 13.6$ eV (the ionization potential of hydrogen) is:

$$\ell_{h\nu} = \frac{1}{n_{\rm H}\, \sigma_{\rm H}} \simeq \frac{0.05}{n_{\rm H}} \quad \text{pc}, \tag{5.54}$$

in which $\sigma_{\rm H} \simeq 6 \times 10^{-18}$ cm^2 is an effective mean of the ionization cross-section of hydrogen. This mean free path is very small so that regions of neutral hydrogen and ionized hydrogen are sharply separated spatially. This is because the transition region between H I and H II regions cannot be much larger than $0.05/n_{\rm H}$ pc.

Interaction between gas and a magnetic field

Interstellar space is permeated by a weak magnetic field with a field strength of typically $B \simeq 6\, 10^{-6}$ Gauss. The magnetic energy density is $B^2/8\pi \simeq 1.4\, 10^{-12}$ erg cm^{-3} or 0.9 eV cm^{-3} , which is comparable to the energy density of the interstellar radiation field.

In a (partially) ionized gas the charged particles move in circular orbits around the magnetic field lines. The orbital angular frequency (also known as cyclotron frequency) is given by:

$$\omega_B = \frac{ZeB}{m}, \tag{5.55}$$

where Z is the charge, in units of the charge e of an electron, and m the mass of the particle. If we take v as the speed of the particle perpendicular to the magnetic field direction we find that $r_B = v/\omega_B$ is the radius of the orbit that the particle makes around the magnetic field line, known as the gyro radius or Larmor radius.

In an H I region which consists mainly of neutral hydrogen, some ionized particles still exist. In particular C atoms (that have an ionization potential of 11.3 eV) can be ionized by photons from the diffuse interstellar radiation field (with energies less than 13.6 eV). For an interstellar magnetic field strength of $B \simeq 6 \times 10^{-6}$ Gauss a C II ion ($m = 12 m_{\rm H}$ en $Z = 1$) has a cyclotron frequency of $\omega_B \simeq 4 \times 10^{-3}$ or, in other terms, an orbital period of 1500 s, much smaller than most other timescales that are relevant in the ISM. For an electron ($Z = 1$) we find $\omega_B \simeq 80$ s^{-1}, so an even smaller orbital period of 0.1 s. A dust particle with mass 10^{-14} g and a charge $Z \simeq 20$ has $\omega_B \simeq 2 \times 10^{-10}$ s^{-1}, so that an orbital period is about 1000 yrs. Even this timescale is much smaller than most other timescales in the ISM. We can conclude

that ions, electrons and charged dust particles are well coupled to the interstellar magnetic field.

As an exercise, calculate the size of the Larmor radius for the given examples. Assume a thermal speed (Equation 2.65) corresponding to a temperature of T = 70 K.

Interaction between gas and cosmic rays

Interstellar space also contains very energetic particles called *cosmic rays*, consisting of protons, α particles (nuclei of He atoms), electrons and nuclei of other elements. The energy density of cosmic rays is $\sim 2 \times 10^{-12}$ erg cm^{-3} or 1.4 eV cm^{-3}, so of the same order of magnitude as the magnetic and radiation energy densities.

Collisions between cosmic rays and the nuclei of H-atoms in the gas lead to the formation of π^0 mesons that decay into two very energetic γ photons with each an energy of $\sim$ 70 MeV

$$\mathrm{H} + \mathrm{cosmic\ ray} \rightarrow \pi^0 \rightarrow 2\gamma. \tag{5.56}$$

This is why gamma-ray radiation can be observed from gas clouds in the plane of the galaxy.

Hydrogen atoms can also be ionized through interactions with cosmic rays. The cross-section for H-ionization increases with decreasing cosmic ray energy. Let us consider a cosmic-ray proton with an energy of 100 MeV. The average ionization frequency of neutral hydrogen for such a proton is $\zeta \simeq 10^{-17}$ s^{-1}, from which we can derive $t_{\mathrm{H-cr}} = \zeta^{-1} \simeq 3\ 10^9$ years. So we find that a hydrogen atom on average is ionized once every $3\,10^9$ yrs by an energetic cosmic ray particle, so about three times in the lifetime of the galaxy. The energy that is transferred to the gas (in the form of thermal energy) is of the order of $\sim$ 3 eV per event. Despite its infrequent occurrence the ionization of hydrogen through cosmic rays can contribute to the heating of interstellar gas. For instance, in the dense cores of molecular clouds or in dense parts of proto-planetary disks the interstellar radiation field cannot penetrate, and ionizations of hydrogen are dominated by cosmic rays.

Exercise 5.1

This could be a nice exam question. The 'very local' interstellar medium has $n_{\rm H} \simeq 0.22\,{\rm cm}^{-3}$ (Lallemont et al. 2004, A&A 426, 875; Slavin & Frisch 2007, Sp.Sci.Revs. 130, 409). The sun is moving at $v_{\rm w} = 26 \pm 1\,{\rm km\,s}^{-1}$ relative to this local gas (Möbius et al. 2004, A&A 426, 897). Suppose that this gas has $n({\rm He})/n({\rm H}) = 0.1$, and contains dust particles with total mass equal to 0.5% of the mass of the gas. Assume these particles have radius $a = 0.1\,\mu$m and specific density $\rho_{\rm s} = 2\,{\rm gr\,cm}^{-3}$. We wish to design a spacecraft to collect them for study. With how large a collecting area A should we equip this spacecraft in order to have an expected collection rate of 1 interstellar grain per hour?

Neglect the motion of the spacecraft relative to the sun, and assume that the interstellar grains are unaffected by solar gravity, radiation pressure, the solar wind, and the interplanetary magnetic field.

Exercise 5.2

a) At time $t = 0$ the population of an excited level is $n_u(0)$. How does the population of this level decrease over time if *only* spontaneous emission to a lower level l occurs and the Einstein coefficient of this process is A_{ul}?

b) Show that the *lifetime* (in sec) of the particle in state u is given by

$$< t >= \frac{1}{A_{ul}} \tag{5.57}$$

Exercise 5.3

Consider a photon of frequency ν entering a slab of homogeneous material containing two-level atoms with excitation temperature $T_{\rm exc}$. At the frequency of the photon, let the optical depth of the slab be $\tau = \tau_{\rm abs} - \tau_{\rm stim}$, i.e. to consist of an absorption component and a stimulated emission component (with a minus sign in front as this component is adding photons to a beam of light).

a) Let $P_{\rm abs}$ be the probability that the original photon will undergo absorption before exiting from the cloud (we do not care about re-emission; once absorbed it is gone from our exercise). Give an expression for $P_{\rm abs}$ in terms of τ and $x = h\nu/kT_{\rm exc}$.

b) Consider a photon that crossed the slab without being absorbed. Let $P_{\rm stim}$ be the probability that the incident photon will stimulate emission of one or more photons. Give an expression for $P_{\rm stim}$ in terms of τ and x.

Designation i	E_i [K]	g_i
$^3P_2$	0.0	5.0
$^3P_1$	227.7	3.0
$^3P_0$	326.6	1.0

Designation $i \to j$	A_{ij} [s^{-1}]	q_{ij} [$cm^3\ s^{-1}$]
$^3P_1 - {}^3P_2$	8.9×10^{-5}	1.5×10^{-10}
$^3P_0 - {}^3P_2$	1.3×10^{-10}	2.4×10^{-10}
$^3P_0 - {}^3P_1$	1.8×10^{-5}	2.1×10^{-12}

Table 5.1: *Left: level designations, energies, statistical weights of hyperfine structure of O I $2p^4\ ^3P$. Right: level designations of the forbidden hyperfine transition, Einstein A coefficient, collision rate coefficient for collisions with molecular hydrogen at 100 K.*

Exercise 5.4

The fine structure lines of neutral oxygen arise from the fact that the ground electronic state splits up into three fine structure levels. The lines are forbidden (see 3.3). Relevant atomic data for the lines are provided in table 5.1.

a) Make a sketch of the energy level structure of the oxygen ground electronic state with the three line transitions and calculate their wavelength.

b) List the processes that impact the level population numbers of these fine structure states.

c) Calculate the critical density of the [O I] 63 μm line transition and compare this value to a typical molecular cloud (use e.g. Table 1.1). Do you expect the level populations to be in Local Thermodynamic Equilibrium in that environment?

d) Predict the LTE total intensity in erg $cm^{-2}\ s^{-1}\ sr^{-1}$ of the [O I] 63 μm line from an optically thin slab of gas with an oxygen column density $N(O) = 10^{12}$ cm^{-2} and constant temperature of 100 K. (Tip: consult Eq. 5.37).

Exercise 5.5

a) Use Eq. 5.49 to estimate the mean temperature of Earth.

b) During its main sequence lifetime the sun will increase its luminosity and radius, and decrease its surface temperature. What will be the effect of these changes on the mean Earth temperature?

6

The interstellar radiation field

We briefly introduced the interstellar radiation field (ISRF) in the introductory chapter (see Sect. 1.3 and Fig. 1.6), and used a simple representation in Sect. 5.4 to make order of magnitude estimates (see Eqs. 5.45 and 5.46). Let us devote a short chapter to better quantify the ISRF in the solar neighborhood. Near home the ISRF is dominated by six components, graphically shown in Fig. 6.1. Their contribution to the total energy density (Eq. 2.12) is given in Table 6.1.

- Galactic synchrotron radiation. The ISM contains relativistic electrons that emit synchrotron emission in the galactic magnetic field; this synchrotron radiation dominates the sky brightness at frequencies $\nu \lesssim 1$ GHz or meter wavelengths and up. This emission is spatially variable. Its intensity is higher near supernovae remnants due in part to electron acceleration associated with the supernova blast wave and in part to increased magnetic field strengths in the shocked gas. In Fig. 6.1 the synchrotron intensity is approximated by

$$(\nu\, u_\nu)_{\rm syn} \approx 3.05 \times 10^{-19} \frac{\nu_9^{0.1}}{1 + 0.04\,\nu_9} \qquad {\rm erg\ cm^{-3}}, \tag{6.1}$$

  where $\nu_9 = \nu/10^9$ Hz. The figure too shows that the synchrotron intensity equals the cosmic background radiation intensity at $\sim$1 GHz. The total energy density in synchrotron radiation obtained by integrating Eq. 6.1 over frequency is small (see Table 6.1).

- The cosmic microwave background radiation. The CMB is essentially isotropic and its intensity is given by a Planck curve at a temperature $T = 2.725$ K, peaking at 1.87 mm in frequency space.

- Free-free, free-bound, and bound-bound transitions from a partially / fully ionized plasma (at $\sim 10^4$ K). We do not discuss this in great detail – it is not an important contributor to the ISRF in the solar neighborhood – but mention that the free-free emission coefficient given by

$$\eta_\nu^{\rm ff} = \frac{8e^6}{3c^3} \left(\frac{2\pi}{3k\,m_{\rm e}^3} \right)^{1/2} \overline{g}_{\rm III}(\nu, T)\, \frac{Z_{jk}^2}{T^{1/2}}\, n_{\rm e}\, N_{jk}\, e^{-h\nu/kT}, \tag{6.2}$$

Component			u [erg cm^{-3}]	u [eV]
Radio synchrotron (Eq. 6.1)			2.7×10^{-18}	1.7×10^{-6}
CMB, $T = 2.725$ K			4.19×10^{-13}	0.262
Dust emission			5.0×10^{-13}	0.312
Free-free, free-bound, two-photon			4.5×10^{-15}	2.8×10^{-3}
Starlight:	$T_1 = 3000$ K, $W_1 = 7 \times 10^{-13}$	4.29×10^{-13}		
	$T_2 = 4000$ K, $W_2 = 1.6 \times 10^{-13}$	3.19×10^{-13}		
	$T_3 = 7500$ K, $W_3 = 1 \times 10^{-14}$	7.11×10^{-14}		
Starlight total			1.05×10^{-12}	0.655
Hα			8×10^{-16}	5×10^{-4}
Other $\lambda \geq 3648$ Å H lines = 1.1× Hα			9×10^{-16}	5.6×10^{-4}
0.1−2 keV X-rays			1×10^{-17}	6.2×10^{-6}
Total ISRF			2.19×10^{-12}	1.367

Table 6.1: *Components contributing to the total energy density (see Eq. 2.12) of the Interstellar Radiation Field (ISRF). From: Draine (2011).*

where N_{jk} is the density of ion j of element k of net charge Z_{jk} and $\overline{g}_{\rm III}$ is the velocity mean for the Gaunt factor of free-free processes; $\overline{g}_{\rm III}$ is dimensionless and of order unity. In cgs-units the front constants together have the numerical value 5.443×10^{-39}. For an ionized hydrogen gas and $\overline{g}_{\rm III}(\nu, T) = 1$ we thus have

$$\eta_\nu^{\rm ff} \;=\; 5.443 \times 10^{-39}\, T^{-1/2}\, n_{\rm e}\, n_{\rm p}\, e^{-h\nu/kT}. \tag{6.3}$$

Note that $\eta_\nu^{\rm ff}$ is proportional to the emission measure $\mathcal{EM}_{\rm H}$ (see Eq. 8.50).

☐ Far-infrared (FIR) and infrared (IR) emission from dust grains heated by starlight. In the solar neighborhood this emission dominates between $\nu \sim 500$ GHz (or 600 μm) and $\sim 6 \times 10^{13}$ Hz (5 μm). About 2/3th of the power radiated by the dust is at $\lambda > 50\,\mu$m and can be approximated as thermal emission from dust particles at a temperature of $T_{\rm d} \sim 17$ K. The 1/3rd of the power emitted at shorter wavelength is dominated by the vibrational bands of PAH molecules as 3.3, 6.2, 7.7, 8.6, 11.3, and 12.7 μm that have undergone single photon heating (see Sect. 12.4).

☐ Starlight. Within an H I region the radiation at energies above 13.6 eV is negligible because of strong absorption by neutral H and He. Only at $\sim 10^2$ eV this absorption has diminished enough to let soft X-rays through. Within the range 1−13.6 eV most of the photons are starlight. Mathis, Mezger, and Panagia (1983) (MMP83) provide an approximate expression for the local starlight background (see below).

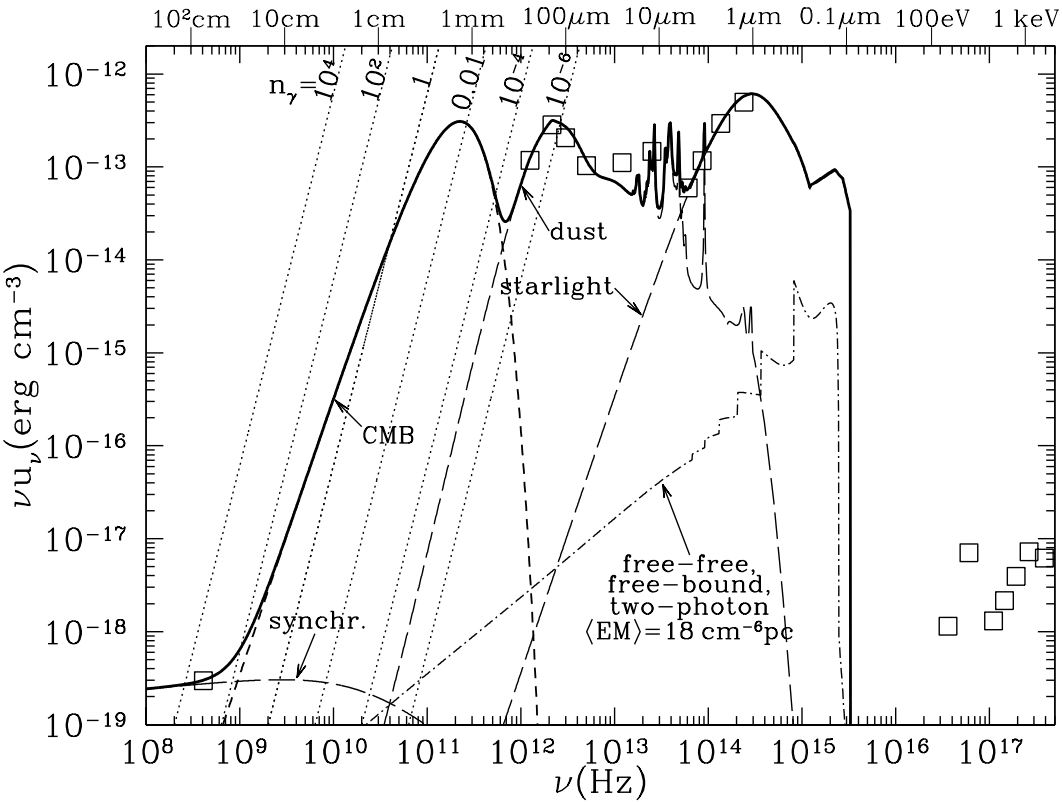


Figure 6.1: *The ISRF in an H*I *cloud in the solar neighborhood. Spectral lines are not included. The solid line is the sum of all components for* $h\nu \leq 13.6$ *eV. Dotted lines are contours of constant photon occupation number* n_γ*. Squares show the measured sky brightness at various frequencies. From: Bruce Draine (2011),* Physics of the interstellar and intergalactic medium.

6.1 The MMP83 prescription for the ISRF

Mathis, Mezger, and Panagia (1983), henceforth MMP83, provide an approximate expression for the local starlight background consisting of the sum of three diluted blackbodies and a piecewise power-law approximation for the range $912\mathring{A} < \lambda < 2450$ Å. The diluted blackbody component is given by

$$\nu\, u_\nu = \sum_{j=1}^{3} \frac{8\pi h\nu^4}{c^3} \frac{W_j}{e^{h\nu/kT_j} - 1} \qquad \text{for } \lambda > 2450\,\text{Å}, \tag{6.4}$$

where the blackbody parameters T_j and W_j are given in Table 6.1. Note that originally, MMP83 adopt $W_1 = 5 \times 10^{-13}$. To improve agreement with the COBE-DIRBY photometry

data, W_1 is increased to 7×10^{-13}. It is this new estimate that is plotted in Fig. 6.1.

The radiation field between 5.04 eV (2460 Å) and 13.6 eV (912 Å) is of considerable importance in the neutral ISM, because it can (1) photo-excite and photo-dissociate H_2 and other molecules, (2) photo-ionize many heavy elements, and (3) eject photo-electrons from dust grains. However, the intensity of the far-ultraviolet (FUV) radiation field is spatially variable because the O and B stars that are the primary emitters are neither numerous nor spatially distributed, and because of the strong attenuating effect of interstellar dust in the FUV. MMP83 represent the local FUV background by three power-law segments. In units of erg cm^{-3} these are

$$\begin{aligned} \nu\, u_\nu &= 2.372 \times 10^{-14}\, (\lambda/\mu\mathrm{m})^{-0.6678} && 1340 - 2460 \text{ Å} \\ &= 6.825 \times 10^{-13}\, (\lambda/\mu\mathrm{m}) && 1100 - 1340 \text{ Å} \\ &= 1.287 \times 10^{-9}\ \ (\lambda/\mu\mathrm{m})^{4.4172} && 912 - 1100 \text{ Å} \end{aligned} \tag{6.5}$$

Early estimates of the background FUV radiation field were made by Habing (1968). He estimated an energy density times frequency $\nu\, u_\nu \sim 4 \times 10^{-14}$ erg cm^{-3} at $\lambda = 1000$ Å (12.4 eV). This is such a well known estimate that it is conventional to reference other estimates of the intensity near 1000 Å to this value. To this end, we define the dimensionless parameter

$$\chi \equiv \frac{(\nu\, u_\nu)_{1000\ \mathring{A}}}{4 \times 10^{-14}\,\mathrm{erg\,cm^{-3}}}. \tag{6.6}$$

The original MMP83 estimate for the ISRF has $\chi = 1.23$. Integration of the spectrum by Habing between 6.0 eV and 13.6 eV gives an energy density $(\nu\, u_\nu)_{6-13.6\,\mathrm{eV}} = 5.29 \times 10^{-14}$ erg cm^{-3}. It is also custom to use this value as a reference. We define

$$G_\circ = \frac{(\nu\, u_\nu)_{6-13.6\,\mathrm{eV}}}{5.29 \times 10^{-14}\,\mathrm{erg\,cm^{-3}}}. \tag{6.7}$$

The original MMP83 estimate for the ISRF has $G_\circ = 1.14$.

The MMP83 estimate of the local starlight intensity (Eqs. 6.4 and 6.5) itself is often used as a reference (see e.g. Sect. 12.4). The starlight component to the ISRF is then described as

$$u_\nu = U\, u_\nu^{\mathrm{MMP83}}, \tag{6.8}$$

where U is a constant.

6.2 Radiation field in a photo-dissociation region (PDR) near a hot star

Consider a luminous massive O star (accompanied by additional lower luminosity stars) exciting an H II region adjacent to a molecular cloud. The Lyman continuum radiation field

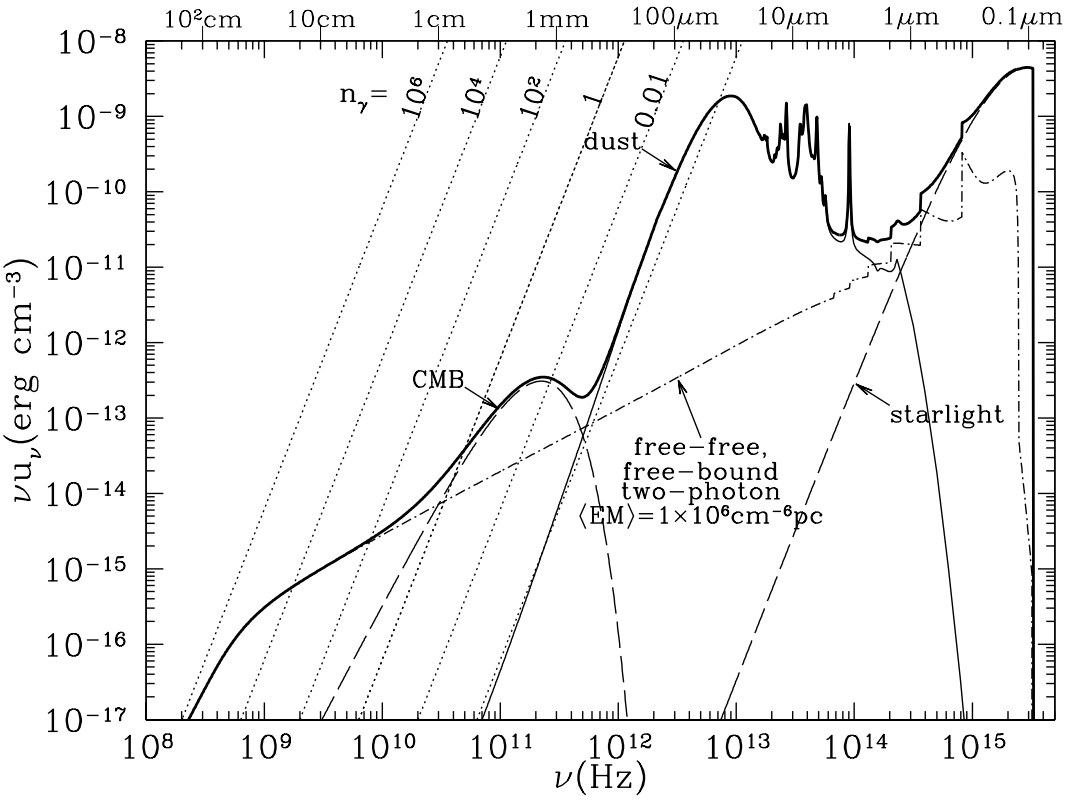


Figure 6.2: *The ISRF in neutral gas adjacent to an H II region with peak* $\mathcal{E}\mathcal{M}_{\rm H} \sim 4 \times 10^6\ cm^{-6}\ pc$. *Starlight is from an O star with* $T = 35,000\ K$ *and dilution factor* $W = 2 \times 10^{-13}$, *which corresponds to a distance of 0.2 pc from a star of spectral type O8 V. Spectral lines, e.g. the hydrogen recombination lines, are not shown. The infrared emission from dust is calculated using a dust model with a starlight heating rate 3000 times the ISRF shown in Fig. 6.1. Dotted lines are contours of constant photon occupation number* n_γ. *From: Bruce Draine (2011),* Physics of the interstellar and intergalactic medium.

(in-between 13.6 eV and $\sim$ 100 eV) will be absorbed within the H II region surrounding the hot star. Radiation below 13.6 eV will arrive at the boundary of the H II region attenuated only by absorption by dust that is able to survive in the H II region. After crossing the ionization front separating the ionized and neutral regions, the photons will enter what is referred to as a *photo-dissociation region* or *photon-dominated region* or PDR. Fig. 6.2 shows the spectrum of the radiation field within the PDR.

The PDR is illuminated by the same CMB and galactic synchrotron radiation as a diffuse H I cloud. In addition, the radiation field includes starlight from the nearby O star at $h\nu < 13.6\,$eV, free-free and line emission from the H II region, and emission from the warm dust in the PDR. From the perspective of the PDR the H II region covers about half of the sky. A bright H II

region, such as the Orion Nebula, has an assumed angle-averaged $\mathcal{E}\mathcal{M}_{\rm H} \approx 1 \times 10^6$ cm^{-6} pc (see Eq. 8.50). It is this value that has been assumed in Fig. 6.2. The starlight that enter the PDR mostly ends up being absorbed by dust, with the energy re-radiated in the IR. The dust emission peaks at several tens of microns or $100-200$ K.

Exercise 6.1

After the Sun, Sirius (α Canis Majoris) is the brightest star in our sky. It is actually a binary; Sirius A and Sirius B. Sirius A is spectral type A1 V, with mass 2.1 $M_\odot$; Sirius B is a (much fainter) white dwarf, with mass 0.98 $M_\odot$. The Sirius system has luminosity $L = 25\,L_\odot$ and is at a distance $d = 2.6$ pc.

a) Give an expression for the total energy density (see eq. 2.12) in terms of luminosity L and distance d.

b) What is the energy density u due to radiation from Sirius alone at the location of the Sun? What fraction of the local starlight background energy density is contributed by Sirius alone? (Consult Table 6.1).

Exercise 6.2

Show, by computation, that indeed the energy density of the Cosmic Microwave Background is $u_{\rm CMB} = 4.19 \times 10^{-13}\,{\rm erg\,cm^{-3}}$, as listed in Table 6.1.

Exercise 6.3

This could be a nice exam question. From the dispersion of radial velocities of galaxies in the Coma cluster and applying the virial theorem, one finds a total mass for the Coma cluster of $3.3 \times 10^{15} M_\odot$. The total mass of all stars in all galaxies of the Coma cluster can be derived from the total UV to infrared luminosity of the entire system and is $1.5 \times 10^{13} M_\odot$. The Coma cluster also contains hot gas, with $T_{\rm gas} = 8.8 \times 10^7$ K. In this exercise we investigate whether this hot gas might account for the difference in total mass and total mass in stars.

Coma is a spherical cluster of galaxies with radius 3 Mpc. The hot gas is optically thin, and we assume that it is distributed homogeneously and composed of fully ionized hydrogen. The temperature of the gas is so high that it effectively emits all its energy at x-ray wavelengths. The measured x-ray luminosity of the gas is $L_{\rm x} = 5 \times 10^{44}$ erg sec^{-1}. We may assume $\overline{g}_{\rm III}(\nu, T) = 1$.

a) Derive a formula that gives the frequency integrated emission of the gas in all directions in erg sec^{-1} cm^{-3}.

b) Derive a formula that gives the electron density as a function of $L_{\rm x}$ and other relevant quantities.

c) Calculate the total mass of the hot gas in solar mass. Recall that the proton mass $m_{\rm p} = 1.66 \times 10^{-24}\,{\rm gr\,cm^{-3}}$. Which fraction of the total mass of the Coma cluster resides in the hot gas component?

d) It appears that the mass in gas and stars can not account for all of the mass in the Coma cluster. What component may be responsible for the missing mass?

7

Cold and Warm Neutral Medium

The distinction between the neutral medium and the ionized medium is to some extend somewhat arbitrary. The neutral medium is best defined by the absence of hydrogen Lyman continuum photons (with $\lambda \leq 912\,\mathring{A}$), so that hydrogen is neutral. These photons apparently can not reach these (therefore neutral) parts of the ISM because of severe interstellar extinction. However, interstellar extinction is considerably smaller at wavelengths longward of the Lyman discontinuity, so that some elements like carbon or the metals as well as the dust grains can be ionized by UV radiation. Cosmic rays ionize a small fraction of all elements even deep inside molecular clouds, and X-rays, when present, can also weakly ionize all elements including hydrogen. The neutral medium has, in fact, a non-zero degree of ionization, which plays a key role in its physics.

As we have seen in Section 1.1, the CNM and WNM together contain most of the mass of the interstellar medium. There are two main observables to study the neutral gas: the 21-cm line of atomic hydrogen and metallic lines. The 21-cm line traces the main constituent of the ISM and allows us to measure its mass(distribution), kinematics and temperature. Allowed transitions of metallic species, many of which are located in the UV and optical part of the spectrum, give the chemical composition and some physical parameters. At far-IR wavelengths one may also detect forbidden fine-structure lines of metal species. These are important coolants of the ISM in media that have temperatures up to $\sim 100\,\mathrm{K}$.

7.1 The 21-cm line of atomic hydrogen

The 21 cm or 1.420 GHz line of hydrogen may be observed in emission or in absorption. Combining equations (5.38) and (5.22), we may write for the line extinction coefficient

$$\chi_\nu^\ell = \frac{c^2}{8\pi\nu_{lu}^2}\frac{g_u}{g_l}A_{ul}\left[1-\frac{g_l}{g_u}\frac{n_u}{n_l}\right] n_l\,\phi_\nu \tag{7.1}$$

$$= \frac{c^2}{8\pi\nu_{lu}^2}\frac{g_u}{g_l}A_{ul}\left[1-e^{-h\nu_{lu}/kT}\right] n_l\,\phi_\nu, \tag{7.2}$$

where the latter equation assumes that the 21-cm line is in LTE (see Sect. 5.2 and Eq. 2.67). In case LTE is not valid, we may still preserve the latter expression replacing T for $T_{\rm ext}$ (see eq. 2.79). As $h\nu_{lu}/kT \ll 1$ (i.e. the Rayleigh-Jeans approximation) we may simplify this expression to

$$\chi_\nu^\ell = \frac{c^2 h}{8\pi k\,\nu_{lu}} \frac{g_u}{g_l} A_{ul}\,\frac{1}{T}\,n_l\,\phi_\nu = 1.040\times10^{-14}\,\frac{1}{T}\,n_l\,\phi_\nu. \tag{7.3}$$

In the neutral medium most of the hydrogen atoms are in the ground state, hence in the two hyperfine sub-levels, because the level immediately above is at an energy of 10 eV. As the two hyperfine sub-levels have $g_u = 3$ and $g_l = 1$ (see figure 3.4) one finds that in LTE the total H I density is given by

$$n_{\rm H} = n_l + n_u = n_l + 3n_l = 4n_l, \tag{7.4}$$

again realizing that the Rayleigh Jeans limit holds. So, $n_l = n_{\rm H}/4$. Using Eq. 2.40 we may derive

$$\tau_\nu = 2.601\times10^{-15}\,\frac{1}{T}\,N({\rm H\,I})\,\phi_\nu, \tag{7.5}$$

where

$$N({\rm H\,I}) = \int_0^D n_{\rm H}(s)\,ds \tag{7.6}$$

is the *column density* of hydrogen atoms in atoms cm^{-2}, i.e. the number of atoms in a column with unit cross-section in the line of sight (see also eq. 4.23). For the 21-cm line line broadening is only caused by the Doppler effect, its natural width being extremely narrow since the lifetime of the upper level is only limited by collisions which are rare. We may thus adopt a Doppler profile for ϕ_ν. This gives

$$\phi_\nu = \phi(\nu-\nu_{lu}) = \phi(\Delta\nu) = \frac{1}{\sqrt{\pi}\,\Delta\nu_{\rm D}}\exp\left[-\left(\Delta\nu/\Delta\nu_{\rm D}\right)^2\right], \tag{7.7}$$

where $\Delta\nu_{\rm D}$ is the *Doppler width* defined as

$$\Delta\nu_{\rm D} \equiv \frac{v_{\rm D}}{c}\nu_{lu} = \frac{\nu_{lu}}{c}\left(\frac{2kT}{m}\right)^{1/2}, \tag{7.8}$$

and $v_{\rm D}$ is the Doppler velocity. It is often convenient to express the line profile as a function of radial velocity rather than of frequency. These are related through $\phi_\nu\,d\nu = \phi_v\,dv$. Using $\Delta v = v - v_\circ = c\,(\nu-\nu_{lu})/\nu_{lu}$, hence $dv/d\nu = c/\nu_{lu}$, this yields

$$\phi_v = \phi(v-v_\circ) = \phi(\Delta v) = \frac{1}{\sqrt{\pi}\,v_{\rm D}}\exp\left[-\left(\Delta v/v_{\rm D}\right)^2\right]. \tag{7.9}$$

For the optical depth in v space we have

$$\tau_v = 2.601\times10^{-15}\,\frac{1}{T}\,N({\rm H\,I})\,\frac{c}{\nu_{lu}}\phi_v = 5.491\times10^{-14}\,\frac{1}{T}\,N({\rm H\,I})\,\phi_v \tag{7.10}$$

as for the 21-cm line $c/\nu_{lu} = \lambda_{lu} = 21.11$ cm (of course). We may multiply both sides with T and integrate over the entire line. As the profile function is normalized to unity (see eq. 5.18) we obtain

$$N(\mathrm{H\,I}) = 1.821 \times 10^{13} \int_{\mathrm{line}} T\,\tau_v\,dv, \tag{7.11}$$

where the front constant has units $(\mathrm{K\,cm\,s^{-1}})^{-1}$. For those that prefer velocities in $\mathrm{km\,s^{-1}}$, the front constant becomes 1.821×10^{18} $(\mathrm{K\,km\,s^{-1}})^{-1}$.

Radiative transfer in the 21-cm line

For a spatially resolved homogenous medium we may use solution (2.78) that conveniently expresses the observed specific intensity in the brightness temperature $T_\mathrm{b}(0)$. By adding and subtracting the background specific intensity similar as to what we did in Eq. 2.55, we have

$$\Delta T_\mathrm{b}(v) = (T - T_\mathrm{bg})\,\left[1 - e^{-\tau_v}\right], \tag{7.12}$$

where $\Delta T_\mathrm{b}(v) = T_\mathrm{b}^\mathrm{obs}(v) - T_\mathrm{bg}$ is the brightness temperature of the line profile at velocity v above the background continuum or baseline level T_bg. For an emission line $\Delta T_\mathrm{b} > 0$ and for an absorption line $\Delta T_\mathrm{b} < 0$.

Let us consider the situation where $T \gg T_\mathrm{bg}$. We may then drop T_bg in Eq. 7.12 and use the resulting equation to eliminate T from Eq. 7.11 such that

$$N(\mathrm{H\,I}) = 1.821 \times 10^{18} \int_{\mathrm{line}} \frac{\Delta T_\mathrm{b}(v)\,\tau_v}{1 - e^{-\tau_v}}\,dv, \tag{7.13}$$

where the velocity is now in $\mathrm{km\,s^{-1}}$. Interestingly, if the medium is optically thin the column density is directly given by a measurement of the spectral line integrated brightness temperature, i.e. it is no longer a function of the kinetic temperature of the medium and the optical depth. We have

$$N(\mathrm{H\,I}) = 1.821 \times 10^{18} \int_{\mathrm{line}} \Delta T_\mathrm{b}(v)\,dv \simeq 3 \times 10^{18}\,\Delta T_\mathrm{b}(v=0)\,v_\mathrm{D}, \tag{7.14}$$

where in the last expression we have approximated the integral over velocity by the ΔT_b at line center times the *full width at half maximum* or FWHM of the peak. For a Doppler profile FWHM = $2\sqrt{\ln 2}\,v_\mathrm{D}$.

Figure 7.1 shows an all-sky map of the H I 21-cm integrated line intensity, converted to $N(\mathrm{H\,I})$, assuming self-absorption to be negligible (so using Eq. 7.14). A map like this allows to study, among others, the thickness of the galactic disk in neutral hydrogen. On the Southern left-side quadrant of the image one can identify the Magellanic Clouds, the SMC on the right and the LMC on the left. Notice that the two dwarf galaxies are conspicuously connected by a bridge of H I gas. Indeed, this feature is called 'The Bridge'.

Let us assume that we have mapped a part of the sky centered on a distant galaxy. We may compute the mass in neutral hydrogen in this galaxy by integrating Eq. 7.14 over the total

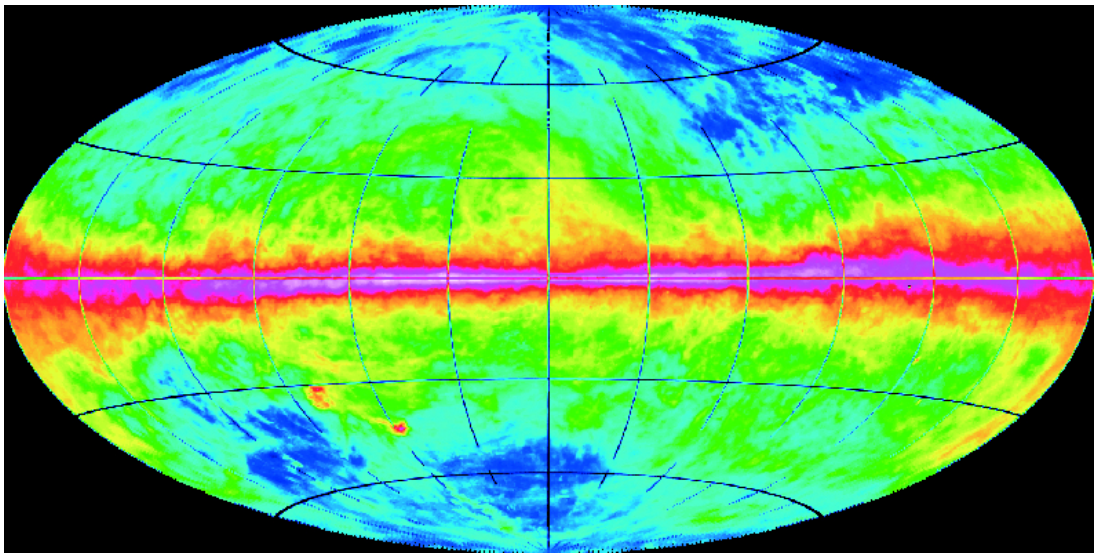


Figure 7.1: *H I emission integrated over the velocity range $-400 < v < +400\,\mathrm{km\,s^{-1}}$ in the Leiden-Argentine-Bonn (LAB) dataset, shown in Hammer-Aitoff projection centered on the Galactic Anti-Center. The angular resolution is 36-arcmin. The LMC and SMC are visible, with a connecting H I 'bridge'. The integrated emission ($0 < N(\mathrm{H\,I}) < 2 \times 10^{22}$ atoms cm^{-2}, on a logarithmic scale) yields column densities under the assumption of optical transparency. This assumption may be violated at latitudes within about 10° of the Galactic equator. From: Kalberla et al. 2005, A&A 440, 775 (but image horizontally mirrored).*

projected surface dS of the galaxy. This is the topic of Exercise 1, so we refrain from doing it here. One finds

$$M[M_\odot] = 2.354 \times 10^5\, d^2\,[\mathrm{Mpc}] \int_{\mathrm{line}} \mathcal{F}_v\,[\mathrm{Jy}]\, dv[\,\mathrm{km/s}], \tag{7.15}$$

where the mass is in solar masses, the distance in Mpc, and the integral is the measured flux integrated over the line in Jansky.

Main findings from studies of the 21-cm line

The main application of the 21-cm line is the measurement of the mass, the distribution and the kinematics of atomic gas in our Galaxy and in external galaxies. For this, it is generally assumed that the line is optically thin, so that the column density does not depend on the physical temperature (Eq. 7.14). Always be conscious that this is not generally correct, consequently that the H I masses obtained in this way may be lower limits. Despite this limitation, emission and absorption measurements of neutral hydrogen have yielded very important insights and results. These include (see Lequeux 2003):

- The atomic interstellar medium is extremely inhomogeneous. The 21-cm emission is

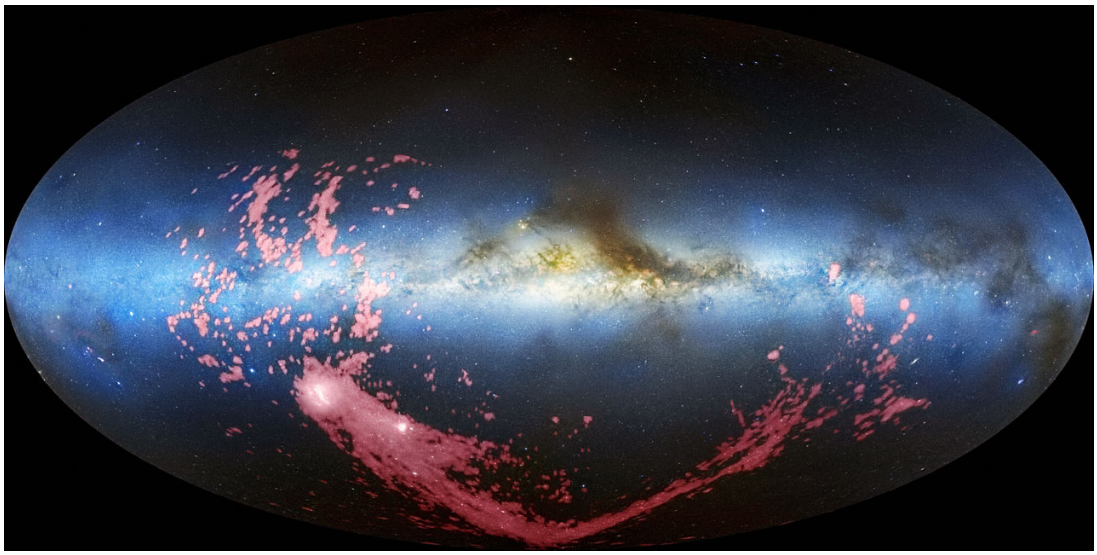


Figure 7.2: *The Magellanic Stream in H* I *(pink) observed with the Green Bank Telescope displayed in Galactic coordinates using an all-sky Hammer-Aitoff projection centered on the Galactic AntiCenter. A model is applied to remove Galactic and non-related extra-galactic emission prior to making this image. The H* I *data is shown together with an optical all-sky image (blue, white, and brown; Mellinger 2009). The Magellanic Bridge between the LMC and SMC and the Leading Arm are clearly visible, as is the extended fragmentation across the Stream. The pink part of the image is similar to Fig. 7.1, but velocity information and an optimized contrast is used to better bring forward the Magellanic Stream. From: Nidever et al. 2010, ApJ 723, 1618 (but image horizontally mirrored).*

dominated by filaments, sheets and shells.

- There are two phases in the atomic interstellar medium: cool H I gas and warm H I gas. They contain about as much matter, however, the warm (several thousands degrees) component has a low density ($0.1-0.5$ cm^{-3}) and is barely seen in absorption. The cool component ($60-100$ K) is denser (tens of atoms cm^{-3}), and dominates absorption. See also Table 1.1. The cool gas is the main contributor to the complex structures mentioned above.

- The warm component forms a thicker galactic disk than does the cool gas. Their respective mean half-thicknesses are 186 pc and 105 pc. Moreover, there seems to exist some neutral gas in the halo, with a high velocity dispersion (of 60 $\mathrm{km\,s^{-1}}$) and a scale height of the order of 4400 pc.

- There exists, at high galactic latitudes neutral gas that falls onto the galactic plane with velocities ranging from a few $\mathrm{km\,s^{-1}}$ to several hundreds of $\mathrm{km\,s^{-1}}$. These *high-velocity clouds* might either be of extragalactic origin, or more probably originate in the hot ionised gas ejected by supernovae and bubbles from the galactic disk (and which forms the Galactic corona) which then falls back onto the disk while cooling and recombining (see Fig. 7.3).

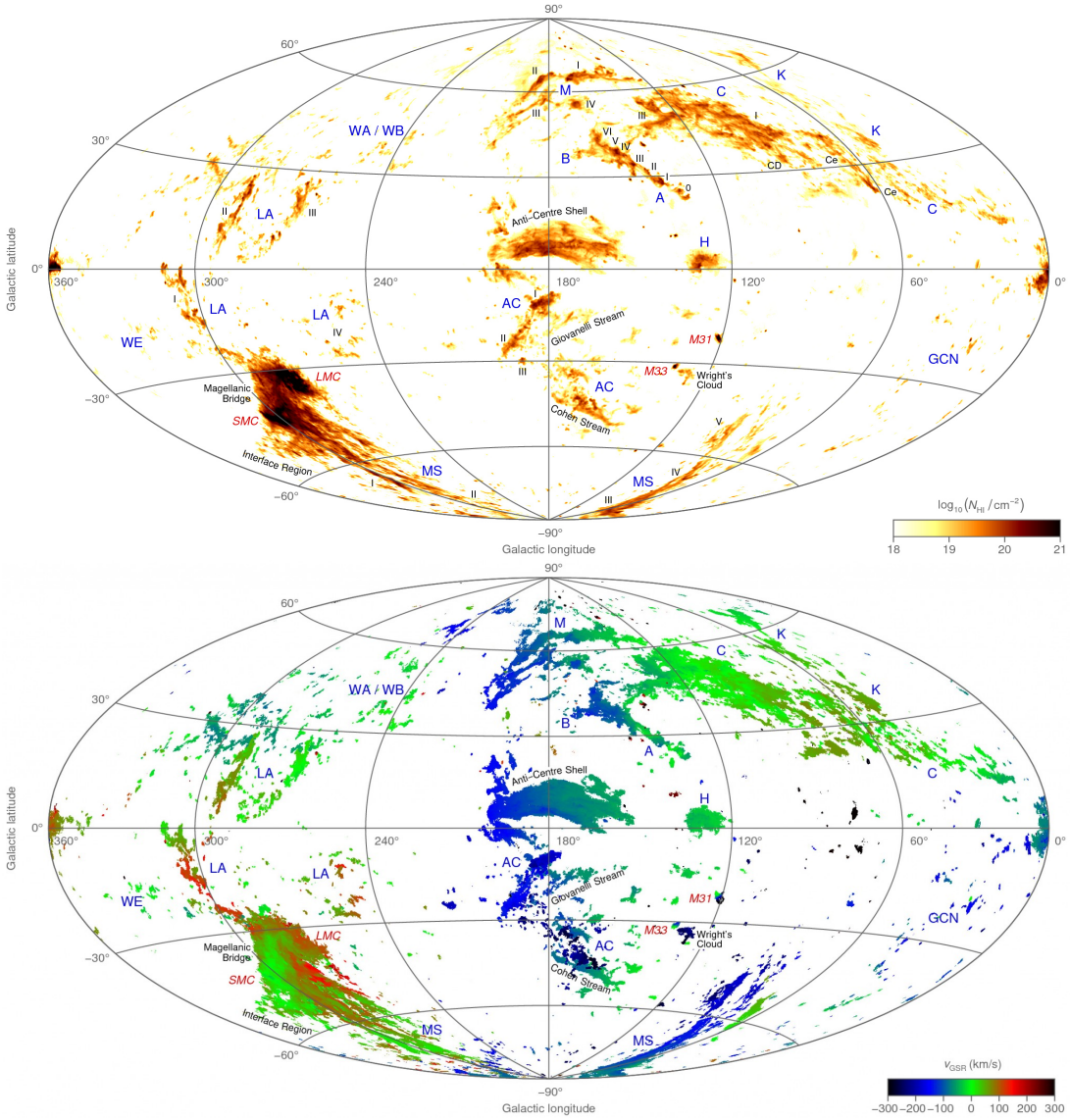


Figure 7.3: *High Velocity Cloud distribution displayed in Galactic coordinates using an all-sky Hammer-Aitoff projection centered on the Galactic AntiCenter. Top: H*I *column density in the range of* $\log_{10}\left(N_{\rm HI}/{\rm cm}^2\right)$ *= 18 to 21 under the assumption that the emission is optically thin and fills the 16.2-arcmin beam. A model is applied to remove Galactic H*I *emission prior to making this image. Bottom: radial velocity in the Galactic standard-of-rest frame, derived from Gaussian fits, in the range of* $v_{\rm GSR} = -300$ *to* $+300 {\rm km\,s}^{-1}$*. Several major HVC complexes (such as Complex C and Complex A at 4−14 kpc and 8−9 kpc distance, and the Magellanic Stream and the Leading Arm) as well as a few notable individual structures and external galaxies (M31, M33) are labelled. Note that a few remaining artefacts and residual stray radiation were manually removed for presentation purposes. From: Westmeier (Tobias) 2018, MNRAS 474, 289.*

- The Magellanic Clouds (MCs) are surrounded by an extended network of gaseous structures. Chief among these is the Magellanic Stream, an interwoven tail of filaments trailing the Clouds in their orbit around the Milky Way. A structure leading the MCs, the Leading Arm, is likely physically related to the Magellanic Stream. The origin of these structures is being investigated. They may be the result from tidal stripping of the MCs by the Milky Way Galaxy (e.g. Gardiner & Noguchi 1996), ram pressure forces exerted by the Galaxy corona (e.g. Moore & Davis 1994), supernova blowout following star formation in the LMC (Nidever et al. 2008), or tidal evolution as a result of a collision between the LMC and SMC (Besla et al. 2010, 2012).

7.2 Interstellar absorption lines

The spectra of stars, safe for those in the direct solar neighborhood, contain many interstellar absorption lines. They differ from stellar lines by being much narrower, the interstellar medium being much cooler than the stellar photospheres where the stellar lines originate. When we target a spectroscopic binary system, we see the photospheric lines of one (in case of a SB1 binary) or both (in case of a SB2 binary) components shift in wavelength while the two sources orbit each other due to the Doppler effect as a result of their orbital motion. Interstellar lines in such a spectrum, however, do not show such a periodic movement. In the visible and near-UV, we observe interstellar lines from atoms (Na, K, Ca), ions (Ca^+, Ti^+), molecules (CN, CH, C_2), and ionized molecules (CH^+). See Fig. 7.4 for an example of a few such lines toward the star ζ Ophiuchi. A very large number of atomic, ionic and molecular lines have been observed in the far-UV. Among the atomic lines the Lyman series of atomic hydrogen lines are conspicuous, and among the molecular lines the many lines of H_2.

These lines contain important information on the chemical composition, physical conditions, and velocity fields in the diffuse interstellar medium. In this section, we will study the line profiles of interstellar lines in more detail. We first define a means to describe the strength of a spectral line. Next, we discuss how one may derive the column densities from such lines.

Describing the line profile

The most complete description of the spectral line is given by its *profile*. The *relative depression* or *absorption depth* of the profile is strickly speaking defined as

$$d_\lambda \equiv 1 - \frac{I_\lambda}{I_\lambda^{\rm c}} \tag{7.16}$$

where $I_\lambda^{\rm c}$ is the continuum intensity at the wavelength λ. In the presence of a spectral line it is by definition impossible to measure the continuum intensity, therefore its value is determined by interpolation of the continuum intensity at both sides of the line profile. In case the absorption depth is positive, the line is an absorption line; in case d_λ is negative, we are dealing with an emission line.

Ion/molecule	λ (Å)	Transition
Na I	(D2) 5,889.95	$3^2S_{1/2} - 3^2P^0_{3/2}$
Na I	(D1) 5,895.92	$3^2S_{1/2} - 3^2P^0_{1/2}$
Ca II	(K) 3,933.66	$4^2S_{1/2} - 4^2P^0_{3/2}$
Ca II	(H) 3,968.47	$4^2S_{1/2} - 4^2P^0_{1/2}$
CN	3,874.61	$B^2\Sigma^+ \leftarrow X^2\Sigma^+(0,0)\,R(0)$
CH	4,300.31	$A^2\Delta \leftarrow X^2\Pi(0,0)\,R_2(1)$
CH^+	4,232.58	$A^1\Pi \leftarrow X^1\Sigma^+(0,0)\,R(0)$
CH^+	3,957.74	$A^1\Pi \leftarrow X^1\Sigma^+(1,0)\,R(0)$

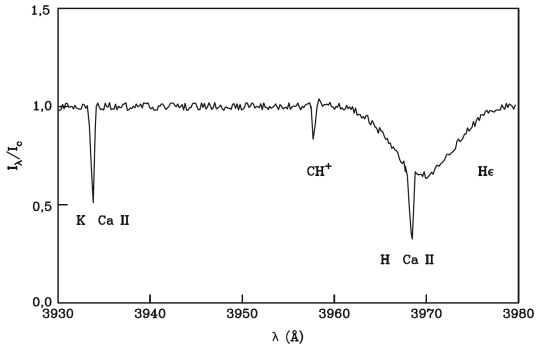


Figure 7.4: *Left panel: a selection of optical absorption lines of interstellar origin. Right panel: A small part of the optical spectrum of ζ Oph, a O9.5 V type stars with $E(B-V)$= 0.32. The Ca* II K *λ3933.66 Å and Ca* II H *λ3933.66 Å as well as the CH^+ λ3957.74 Å interstellar lines can be observed. Note that the Ca* II H *line is positioned inside the* Hϵ *hydrogen line, which is much broader and of stellar origin. Table and figure: Maciel (2013)*

If one can not (or does not want to) measure the specific intensity, one may use a measurement of the flux to describe the absorption depth. In this case

$$D_\lambda \equiv 1 - \frac{\mathcal{F}_\lambda}{\mathcal{F}^{\rm c}_\lambda} \tag{7.17}$$

where $\mathcal{F}^{\rm c}_\lambda$ is the continuum flux at wavelength λ. Also the continuum flux at a wavelength in the line profile can not be measured directly, and therefore it also follows from interpolation of the continuum flux at both sides of the line.

Before the absorption profile can be used as a diagnostic of the interstellar gas it needs to be corrected for instrumental distortions. This distortion, which always leads to degradation of the line profile, is described by the *instrumental profile*. Say a light source emits an emission line of infinitesimal width. The profile of this line can be described by a δ-function. Measurement of this line by an instrument shows a smeared profile (typically a Gaussian profile), of which the sharpness is determined by the quality and/or settings of the spectrograph. A measure of this sharpness is the *spectral resolution* or *resolving power*

$$R \equiv \lambda/\Delta\lambda \tag{7.18}$$

Here $\Delta\lambda$ is the bin-width of the flux measurement. For excellent instruments R can be larger than 100 000.

Equivalent width

The concept of *equivalent width* of a spectral line was developed by Marcel Minnaert (1893-1970)[1]. The equivalent width is the line profile integrated absorption depth, i.e.

$$W_\lambda(\text{line}) \equiv \int_{\text{line}} D_\lambda \, d\lambda = \int_{\text{line}} \left(1 - \frac{\mathcal{F}_\lambda}{\mathcal{F}_\lambda^{\text{c}}}\right) d\lambda \tag{7.19}$$

and is – in case of an absorption line – equivalent to the width of a fully blackened rectangular profile of identical surface area (see figure 7.5). It is custom to measure W_λ in (milli-) Angström or in the velocity unit $\mathrm{km\,s^{-1}}$. In case of an absorption line $\mathcal{F}_\lambda^{\text{c}} \cdot W_\lambda$ is equal to the total continuum energy that is removed by the line. In case of an emission line W_λ will be negative, and $-\mathcal{F}_\lambda^{\text{c}} \cdot W_\lambda$ describes the total energy that is added by the line to the continuum. The equivalent width is a suited measure for the strength of the spectral line, as it is, for instance, much less sensitive to smearing of the profile as a result of the finite resolution of the spectrograph, than is, for instance, the central absorption depth. For an accurate measurement of W_λ it usually suffices to have a spectral resolution of $R \sim 8\,000$ if the signal-to-noise is at least several tens. The equivalent width offers a quantitative measure of the line profile in cases where the flux levels are too low to observe a detailed line profile. Finally, the equivalent width is independent of interstellar extinction (see § 12.2).

If the source is spatially resolved, one can determine the equivalent width from

$$w_\lambda(\text{line}) \equiv \int_{\text{line}} d_\lambda \, d\lambda \tag{7.20}$$

Total line flux

The line profile integrated flux or *total line flux* in $\mathrm{erg\,cm^{-2}\,sec^{-1}}$ is

$$\mathcal{F}(\text{line}) \equiv \int_{\text{line}} (\mathcal{F}_\nu - \mathcal{F}_\nu^{\text{c}}) \, d\nu = \int_{\text{line}} (\mathcal{F}_\lambda - \mathcal{F}_\lambda^{\text{c}}) \, d\lambda \tag{7.21}$$

The total line flux (measured at distance d) is especially important in studies of emission lines, and therefore is defined such that a positive value results if $\mathcal{F}_\nu > \mathcal{F}_\nu^{\text{c}}$. One obtains $\mathcal{F}(\text{line}) = -\mathcal{F}_\lambda^{\text{c}} \cdot W_\lambda = -\mathcal{F}_\nu^{\text{c}} \cdot W_\nu$. In, for instance, planetary nebulae the continuum flux can be so low that it can not be measured (accurately), i.e. $\mathcal{F}_\nu \gg \mathcal{F}_\nu^{\text{c}}$. The continuum contribution can then simply be omitted when computing the total line flux. Note that the total line flux – in contrast with the equivalent width – is dependent on the amount of interstellar extinction.

Related to the total line flux is the *luminosity in the line*, in $\mathrm{erg\,sec^{-1}}$, i.e.

$$L(\text{line}) = 4\pi d^2 \, \mathcal{F}(\text{line}) \tag{7.22}$$

[1]Born in Ghent, Belgium, Minnaert worked at the Utrecht astronomical observatory Sonnenborgh from shortly after World War I. From 1937 until 1962 he was the director of the observatory.

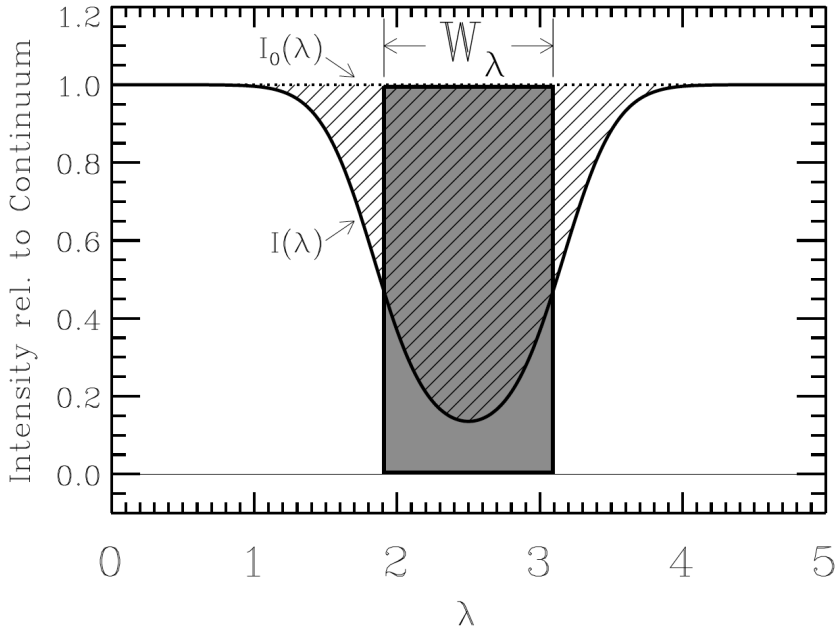


Figure 7.5: *Schematic representation of the flux behaviour of a spectral line (i.e. the profile). The equivalent width W_λ is equal to the surface area of the spectral line divided by the continuum flux. Figure: Edward Jenkins.*

7.3 Curve of growth

We consider a simple model for interstellar absorption in the line of sight towards a distant star. The star is responsible for the continuum light and emits at a brightness temperature T_b. In front of the star is a homogeneous layer where the interstellar absorption line is formed, and that emits according to a Planck function at temperature T_L. We will assume that $T_\mathrm{L} < T_\mathrm{b}$, such that an absorption line is formed. Note that if the temperature in the line forming layer would be higher than T_b – implying an emission line – the following discussion would remain valid.

Though mostly we have been considering frequency space, we will now switch to wavelength space as this is quite generally done in absorption line studies. The emerging intensity follows from eq. (2.52) and is equal to

$$I_\lambda = B_\lambda(T_\mathrm{b})\, e^{-\tau_\lambda} + B_\lambda(T_\mathrm{L}) \left[1 - e^{-\tau_\lambda}\right] \tag{7.23}$$

where τ_λ is the optical depth in the spectral line. If we neglect the contribution of stimulated emission in the line extinction coefficient, we have

$$\chi_\lambda = \alpha_{lu}(\lambda)\, n_l = \frac{\pi e^2}{m_\mathrm{e} c} \frac{\lambda_{lu}^2}{c} f_{lu}\, \phi(\lambda)\, n_l \tag{7.24}$$

where $(\pi e^2/m_{\rm e}c) = 0.02654\,{\rm cm}^2\,{\rm hz}$. f_{lu} is the oscillator strength. It is related to the Einstein coefficients in the following way

$$\frac{\pi e^2}{m_{\rm e}c}\,f_{lu} = \frac{h\nu_{lu}}{4\pi}\,B_{lu}. \tag{7.25}$$

For the line profile function $\phi(\lambda)$ we consider a Voigt profile, which is the convolution of a Doppler profile and a Lorentz profile. This is a quite general expression for the profile shape, capturing thermal and turbulent broadening processes (described by a Doppler profile) and natural broadening and collisional (or pressure) broadening (described by a Lorentz profile). Let us briefly summarize the essentials of these profiles. Doppler profile Eq. 7.7 in wavelength space is given by

$$\phi_\lambda = \phi(\lambda - \lambda_{lu}) = \phi(\Delta\lambda) = \frac{1}{\sqrt{\pi}\,\Delta\lambda_{\rm D}}\exp\left[-\left(\Delta\lambda/\Delta\lambda_{\rm D}\right)^2\right], \tag{7.26}$$

where the Doppler width $\Delta\lambda_{\rm D} = (\lambda_{lu}^2/c)\,\Delta\nu_{\rm D}$. The Lorentz profile in wavelength space is

$$\phi_\lambda = \frac{1}{\pi}\,\frac{\Delta\lambda_{\rm L}}{\Delta\lambda^2 + \Delta\lambda_{\rm L}^2}, \tag{7.27}$$

$$\tag{7.28}$$

where $\Delta_{\rm L} = \lambda_{lu}^2\,\gamma/4\pi c$ is the Lorentz width and γ the damping constant. The convolution of these two profiles yields the Voigt profile

$$\phi_\lambda = \frac{\lambda_{lu}^2}{c}\,\frac{1}{\sqrt{\pi}\Delta\lambda_{\rm D}}\,H(a,x), \tag{7.29}$$

where $H(a,x)$ is the *Hjerting function*. The Hjerting function is not normalized – unlike the Doppler and Lorentz profile – but has a surface $\sqrt{\pi}$ in units of x (see below). Examples of the Hjerting function are given in Fig. 7.6. We have

$$H(a,x) \equiv \frac{a}{\pi}\int_{-\infty}^{+\infty}\frac{\exp(-y^2)}{(x-y)^2 + a^2}\,dy, \tag{7.30}$$

where $x = \Delta\lambda/\Delta\lambda_{\rm D}$ is the distance from line center in units of the Doppler width and $a = \Delta\lambda_{\rm L}/\Delta\lambda_{\rm D}$ is the ratio of the natural width to the Doppler width. In most astrophysical circumstances $a \ll 1$, in which case the Hjerting function can be approximated quite accurately by the sum of a Gaussian core and the damping wings of the Lorentz profile, i.e.

$$H(a,x) \simeq \exp(-x^2) + \frac{a}{\sqrt{\pi}x^2}. \tag{7.31}$$

Note that $H(a,0) \simeq 1 - a/\sqrt{(\pi)}$.

If we define the integrated *column density* in ${\rm cm}^{-2}$ of particles in level l that are in the line of sight similar to Eq. 7.6, i.e.

$$N_l \equiv \int n_l(s)ds, \tag{7.32}$$

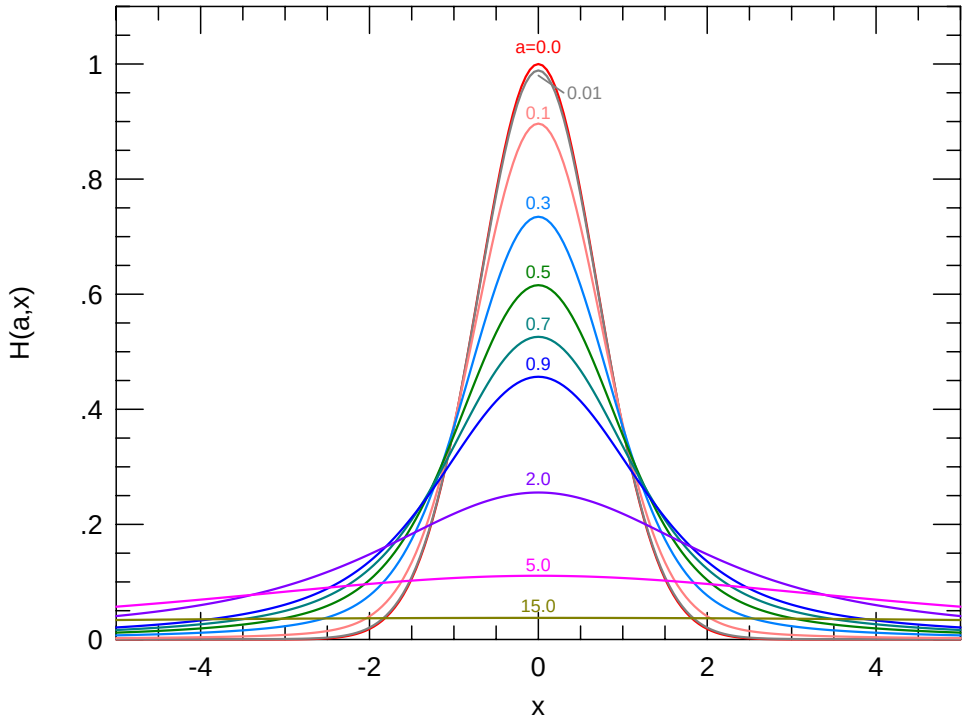


Figure 7.6: *Hjerting functions for different a. For small a and in case of weak spectral lines the central parts of the profile can well be approximated with a Doppler profile. For large a and strong strong lines the wings are characterized by the Lorentz behavior $H(a,x) \propto x^{-2}$. From: Cococubed School of Earth and Space Exploration Arizona State University. See: Cococubed.com.*

we obtain for the line optical depth

$$\tau_\lambda = \alpha_{lu}(\lambda)\, N_l = \frac{\pi e^2}{m_e c}\,\frac{\lambda_{lu}^2}{c}\, f_{lu}\,\frac{1}{\sqrt{\pi}\Delta\lambda_D}\, N_l\, H(a,x) = \frac{\tau_\circ}{H(a,0)}\, H(a,x) \simeq \tau_\circ\, H(a,x), \tag{7.33}$$

where $\tau_\circ$ is the optical depth at line centre. The last equality in Eq. (7.33) follows from $H(a,0) \simeq 1 - a/\sqrt{(}\pi) \simeq 1$ for $a \ll 1$. For the relative depression d_λ we get

$$d_\lambda = 1 - \frac{I_\lambda}{I_\lambda^c} = \frac{B_\lambda(T_b) - B_\lambda(T_L)}{B_\lambda(T_b)}\left(1 - e^{-\tau_\lambda}\right) \equiv d_\lambda^{\max}\left(1 - e^{-\tau_\lambda}\right), \tag{7.34}$$

where $I_\lambda^c = B_\lambda(T_b)$, the continuum specific intensity in the absence of the spectral line. $d_\lambda^{\max}$ is the maximum depression. This yields for the equivalent line width

$$w_\lambda = d_\lambda^{\max} \int_{\text{line}} \left(1 - e^{-\tau_\lambda}\right) d\lambda. \tag{7.35}$$

Figure 7.7 shows how the line profile and equivalent line width depend on optical depth $\tau_\circ$ for a line that has $d_\lambda^{\max} = 1$, i.e. in case the source function in the line forming layer – i.e.

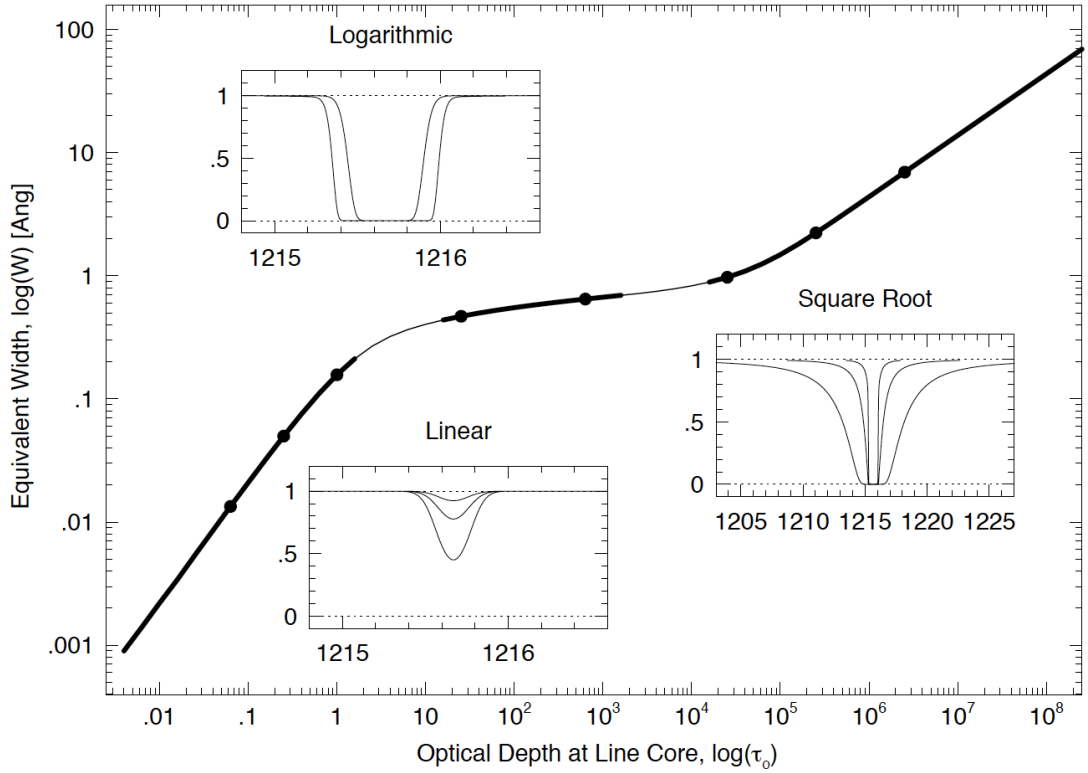


Figure 7.7: *The curve of growth for a Lyα of H I with a Doppler velocity $\xi_{\rm D} = 30\,{\rm km\,s^{-1}}$. The three regimes discussed in the text, the linear (weak lines), flat (saturated lines), and damping part (strong lines) are shown by thicker curves. The approximations describing these regimes are Eq. 7.36, 7.37, and 7.39, respectively. Corresponding line absorption profiles are shown for each regime and their locations on the COG are marked with filled dots. Figure: Chris Churchill.*

the interstellar medium – is negligibly small compared to the continuum source function. The behaviour of the equivalent width may be characterized by three regimes.

Weak lines

First, the regime of weak lines, for which $\tau_\lambda \ll 1$. In this limit the relative depression reduces to $d_\lambda \simeq d_\lambda^{\rm max}\,\tau_\lambda$. For weak lines we may replace the Hjerting function $H(a,x)$ by $\exp[-(\Delta\lambda/\Delta\lambda_{\rm D})^2]$, which has surface $\sqrt{\pi}\Delta\lambda_{\rm D}$. This yields

$$w_\lambda \simeq d_\lambda^{\rm max}\tau_\circ\sqrt{\pi}\Delta\lambda_{\rm D} = \frac{\pi e^2}{m_{\rm e}c}\frac{\lambda_{lu}^2}{c} f_{lu}\, d_\lambda^{\rm max}\, N_l \tag{7.36}$$

The equivalent width increases linearly with column depth N_l. The curve of growth therefore first shows a linear increase of w_λ with line strength. Note that the equivalent width is independent of the profile function.

Saturated lines

For $\tau_\circ > 1$ the core of the line saturates. The intensity at line centre approaches the value $B_\lambda(T_\mathrm{L})$, reflecting the maximum depression d_λ^max. The width of the line wings still increases, however, the corresponding increase in the equivalent width is no longer linear with $\tau_\circ$, but proceeds at a slower pace. It approximately holds that

$$w_\lambda \approx d_\lambda^\mathrm{max}\, Q(\tau_\circ)\, \Delta\lambda_\mathrm{D}. \tag{7.37}$$

For increasing optical depth, $Q(\tau_\circ)$ increases from about 2 to 6 in this *plateau* or *shoulder part* of the curve of growth. An approximate analytical expression for this function is $Q(\tau_\circ) = 2\sqrt{\ln \tau_\circ}$, valid from $\tau_\circ \gtrsim 3$. This regime of the curve of growth is therefore (also) referred to as the *logarithmic part*.

Strong lines

For very strong lines, i.e. lines for which $\tau_\circ \gg 1$, the line core is completely saturated and does not react to a further increase of $\tau_\circ$. However, the far line wings will still have $\tau_\lambda < 1$. For a sufficiently large $\tau_\circ$ both wings will contribute significantly because they are formed in the damping part of the Voigt profile, i.e. where $H(a,x) \simeq a/(\sqrt{\pi}x^2) = (a/\sqrt{\pi})(\Delta\lambda_\mathrm{D}/\Delta\lambda)^2 \sim 1/\Delta\lambda^2$. This decrease with wavelength is much less dramatic than the exponential decay of the Doppler core. In the damping part of $H(a,x)$ we may write

$$\tau_\lambda = \tau_\circ \frac{a}{\sqrt{\pi}x^2} = \tau_\circ \frac{a}{\sqrt{\pi}} \frac{\Delta\lambda_\mathrm{D}^2}{\Delta\lambda^2}. \tag{7.38}$$

Using the transformation $u^2 = 1/\tau_\lambda$, we obtain after substitution in eq. (7.35) for the equivalent width

$$\begin{aligned} w_\lambda &\simeq d_\lambda^\mathrm{max} \Delta\lambda_\mathrm{D} \sqrt{\tau_\circ \frac{a}{\sqrt{\pi}}} \int_\mathrm{line} (1 - e^{-1/u^2})\, du \\ &= d_\lambda^\mathrm{max} \Delta\lambda_\mathrm{D} \sqrt{\tau_\circ a}\, 2\pi^{1/4}. \end{aligned} \tag{7.39}$$

For the last equality we used the standard integral

$$\int_{-\infty}^{+\infty} \left(1 - e^{-1/x^2}\right) dx = 2\sqrt{\pi}. \tag{7.40}$$

Empirical curve of growth and gas-phase elemental abundances in the ISM

Let us consider the line of sight toward ζ Ophiuchi in the constellation Ophiuchus. The visual spectrum of this bright O9.5 V star is well known for its many interstellar absorption lines. Studies have identified at least six distinct H I clouds and one H II cloud in this 112 parsec

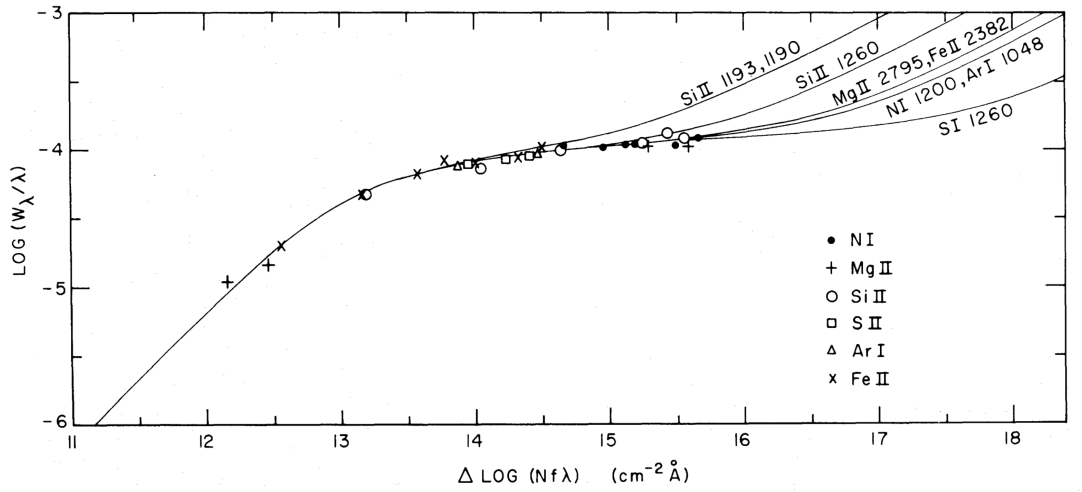


Figure 7.8: *Empirical curve of growth for the dominant ion stages expected in* H I *clouds. The chosen line of sight is toward the O9.5 V star* ζ *Ophiuchi. The turbulent velocity is 6.5* $\mathrm{km\,s^{-1}}$ *and the damping constants appropriate for the lines labeled in the upper right corner. The horizontal scale was labeled to give the column depth for* Fe II *2382 Figure: Donald Morton, 1975, ApJ 197, 85.*

long sight line. ζ Oph is moving through space with a peculiar velocity of 30 $\mathrm{km\,s^{-1}}$ and is a *runaway star* from the Scorpius-Centaurus Association.

The different H I clouds can be identified through their distinct radial velocities. We assume that absorption lines from different elements in the same cloud are broadened in the same way by the Doppler effect, with the same velocity dispersion. This is the case if the Doppler effect is dominated by turbulence rather than by thermal motion, frequently the case for interstellar clouds. In this case the Doppler width is given by (see Eq. 7.8)

$$\Delta\nu_{\mathrm{D}} = \frac{v_{\mathrm{D}}}{c}\nu_{lu} = \frac{\nu_{lu}}{c}\sqrt{v_{\mathrm{th}}^2 + v_{\mathrm{turb}}^2} = \frac{\nu_{lu}}{c}\sqrt{\left(\frac{2kT}{m}\right) + v_{\mathrm{turb}}^2} \simeq \frac{\nu_{lu}}{c}\,v_{\mathrm{turb}} \tag{7.41}$$

hence $\Delta\lambda_{\mathrm{D}} = (\lambda_{lu}^2/c)\,\Delta\nu_{\mathrm{D}} = (\lambda_{lu}/c)\,v_{\mathrm{turb}}$, where v_{turb} is the turbulent velocity.

The Doppler width $\Delta\lambda_{\mathrm{D}}$ is proportional to λ_{lu} and for unsaturated lines $w_\lambda \propto \Delta\lambda_{\mathrm{D}} \propto \lambda_{lu}$ (see Eq. 7.36). As a result, if we plot the *curve of growth* which gives $\log(w_\lambda/\lambda_{lu})$ as a function of $\log(\lambda_{lu}\,N_l\,f_{lu})$, the representative points for the different lines from lower level l are located on a single curve, even if these lines are saturated. Be ware that this is no longer the case for larger column densities, in which case the line width is dominated by the damping wings, because the damping constant is not the same for all lines.

Figure 7.8 shows the empirical curve of growth for ζ Oph. Lines of different elements are shifted such that they all are given as a function of the column depth of Fe II λ2382. To do so requires knowledge of the excitation, ionization, and elemental abundance of the species

involved. In a formal treatment one may split out these quantities in the following way

$$\begin{aligned} N_{ljk} &= \frac{N_{jlk}}{N_{jk}} \frac{N_{jk}}{N_k} \frac{N_k}{N_N} N_N \\ &= \frac{g_{ljk}}{U_{jk}(T_{\rm ext})} e^{-E_{ljk}/kT_{\rm ext}} q_{jk}(T_{\rm ion}) A_k N_N \end{aligned} \tag{7.42}$$

where $N_{ljk} = N_l$ in our previous notation; N_{jk} is the column depth of the ion; N_k is the elemental abundance, and N_N is the column depth of hydrogen. In this notation $q_{jk} = N_{jk}/N_k$ is the ionization fraction. The notation adopts the LTE formalism (see Eq. 2.70). However, the interstellar lines seen in the ISM all originate from the ground level, hence $N_{ljk}/N_{jk} \simeq 1$. For the ionization temperature it is safe to adopt the kinetic temperature derived from the 21-cm line. Note that the ionization equilibrium is a function of temperature and electron density (see e.g. eq. 2.72). For some elements it can be difficult to estimate their total abundance if only one ionization state can be observed. For O, N, and the noble gases this problem does not arise because they are not ionized in the neutral (cold and warm) medium. There is no difficulty either for most of the metals – including iron – and carbon which essentially exist as singly ionized species in this medium. Calcium typically exists as Ca^0 and Ca^+, though Ca^+ is much more abundant. The ionization equation of Calcium (for ionization equations, see section 8.3) can then be used to constrain the electron density $n_{\rm e}$. The electron density in warm clouds is generally found to be lower than 1 electron $\rm cm^{-1}$ (see Table 1.1). Most of these free electrons are in interstellar clouds that are relatively transparant to the UV, and come from the ionization of carbon, the most abundant element with an ionization potential (11.260 eV) lower than that of hydrogen. Once the state of the gas has been established, abundances relative to hydrogen (or, as in Fig. 7.8, relative to iron) may be determined by shifting curves of growth for different elements.

The result of such analysis have already been shown in Fig. 1.5. Most elements are found to be underabundant in the diffuse interstellar medium with respect to the Solar system, young stars and H II regions. Exceptions are S, Zn and P, which do have solar abundances in the warm neutral medium. The missing elements in the ISM are believed to be in dust grains. The underabundance of an element is larger if its condensation temperature is higher, confirming this idea. Condensation must have taken place in circumstellar envelopes, and also the interstellar medium itself. The underabunances are smaller in the warm medium, suggesting that the evaporation of grains, probably as a consequence of shocks, has returned a fraction of the elements to the gas phase.

Exercise 7.1

Prove that Eq. 7.15 follows from integrating Eq. 7.14 over a projected surface area dS. The source is at a very large distance, such that $dS = d^2\, d\omega$, where d is the distance to the source, and $d\mathcal{F}_\nu = I_\nu\, d\omega$, where $d\mathcal{F}_\nu$ is the flux we receive from solid angle $d\omega$.

Exercise 7.2

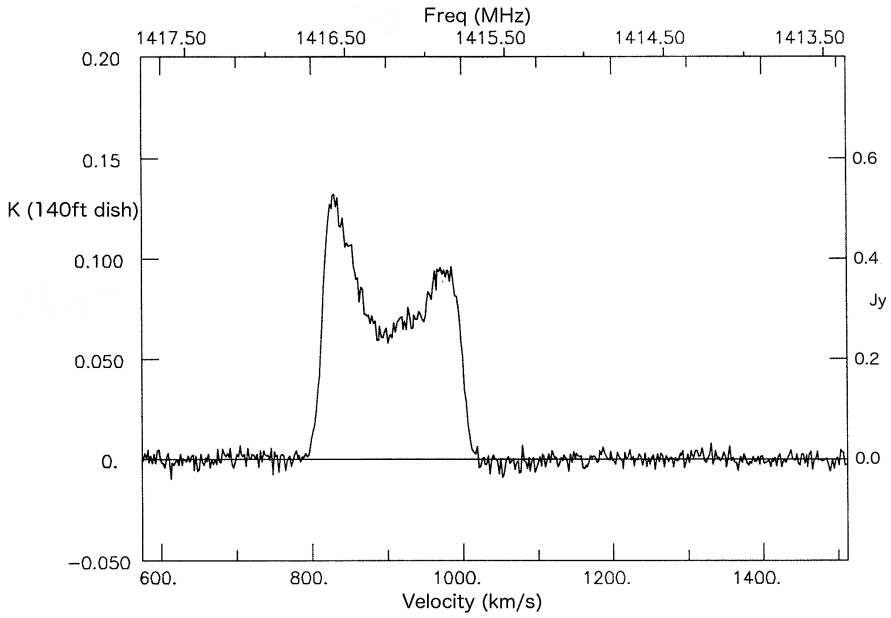


Figure 7.9: *Integrated H I spectrum of the galaxy UGC 11707 obtained with the 140-foot telescope of the* National Radio Astronomy Observatory *(NRAO) in Green Bank, West Virginia. The beam width ∼ 20 arcmin. The line shows the typical two-horned profile of a spiral galaxy. The line center frequency is $\nu = 1416.2$ MHz. The rest wavelength line center frequency is at $\nu_\circ = 1420.4$ MHz or 21 cm. The measured line width is 200 km/s. From: NRAO.*

We have measured the H I 1420.4 MHz (21 cm) line of the spiral galaxy UGC 11707 with the NRAO facility in Green Bank (see Fig. 7.9).

a) Assume the galaxy is at such a large distance that its recession velocity is caused by the uniform Hubble expansion of the universe. Derive the distance to this galaxy. The Hubble constant is 72 km s^{-1} Mpc^{-1}.

b) Assume that the H I emission from this galaxy is optically thin. The mean peak height of the line is at about 0.35 Jy. Compute the H I mass of UGC 11707.

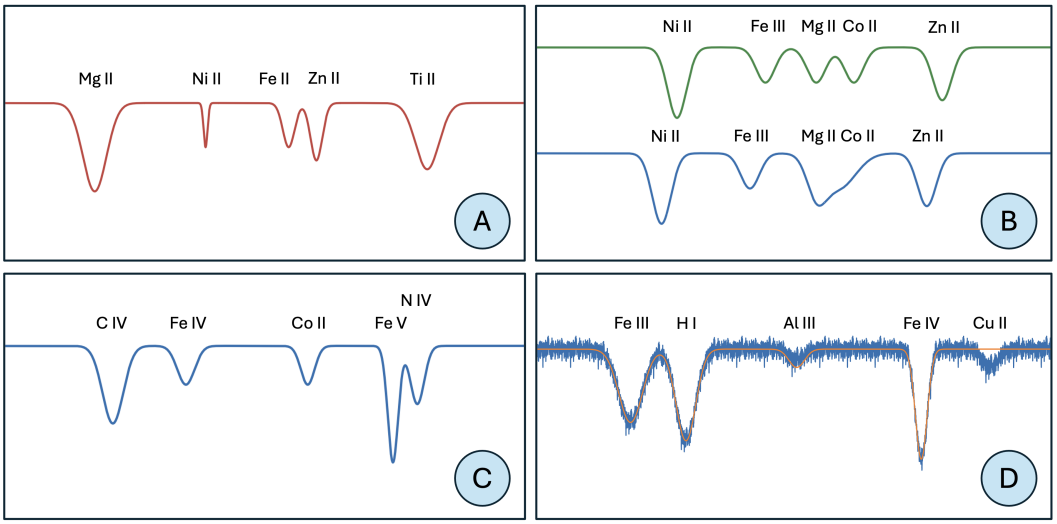


Figure 7.10: *Panels showing a hypothetical spectral window containing stellar and one interstellar line. The spectral resolution of the spectrum in panel A is very high; that of the other panels is so low that the line broadening is dominated by the spectral resolution of the instrument. Hints: In panel A the interstellar cloud in front of the star is cold relative to the stellar temperature. In panel B two observations of the star are shown taken several days apart. In panel D the noisy stellar spectrum is fitted with a model atmosphere.*

Exercise 7.3

In Fig. 7.10, each panel shows a hypothetical spectral window containing stellar and one interstellar line. Identify in each panel the interstellar line and briefly explain your reasoning. The spectral resolution of the spectrum in panel A is very high; that of the other panels is so low that the line broadening is dominated by the spectral resolution of the instrument. Hints: In panel A the interstellar cloud in front of the star is cold relative to the stellar temperature. In panel B two observations of the star are shown taken several days apart. In panel D the noisy stellar spectrum is fitted with a model atmosphere.

Exercise 7.4

Show that the equivalent width in frequency units is

$$W_\nu(\text{line}) \equiv \int_{\text{line}} \left(1 - \frac{\mathcal{F}_\nu}{\mathcal{F}_\nu^c}\right) d\nu \tag{7.43}$$

and that it is related to the equivalent width in wavelength units as $W_\nu = c/\lambda^2 W_\lambda$.

8

H II regions - part I: structure and dynamical state

In this and the next chapter we study the physics of ionized regions of hydrogen, so-called H II regions. In this chapter we first summarize the types of H II regions. Then we address the ionization structure and dynamics of these clouds. In then next chapter we study the energetics of H II regions. The latter requires us to study the heating and cooling processes of interstellar gas. Finally, we discuss how one obtains key properties of H II regions, i.e. its temperature and density.

8.1 Introduction

In the 19th century an animated debate ensued on the question whether the diffuse objects seen through telescopes, loosely termed 'nebulae', were in fact unresolved stellar associations or gas clouds. The answer came when these objects were studied using spectroscopy. The small 'white nebulae', those located primarily outside of the plane of the Milky Way, featured spectra typical for that of stars. These are stellar associations or star clusters. Other nebulae showed a pronounced line spectrum, sometimes accompanied by an underlying continuum that did not appear to be of stellar origin. These are the gas clouds, including planetary nebulae, H II regions, and supernova remnants.

The spectrum of the brightest of these nebulae, the Orion nebula (see Fig. 8.1), was observed in 1863 by William Huggins. Soon after, Balmer lines were recognized in this and other nebulae. After it was discovered in the Sun, also helium was found. Surprisingly enough, the identification of the by far strongest lines in many nebular spectra – in green light at 4959 and 5007 Å – and those of other strong transitions, turned out to be a challenge. Some even speculated that these unidentified lines were produced by a new and hypothetical element 'nebulium'. Progress in the understanding of atomic structure eventually led to the identification of these lines as forbidden transitions of O III. Other strong nebular lines were found to be of a similar

Figure 8.1: *The Orion Nebula (also known as Messier 42, M42, or NGC 1976) at a distance of ∼400 pc is the closest region of massive star formation to Earth. The dimensions of this image are* 65×60 *arcmins, corresponding to a physical scale of* 7.5×7.0 *pc. The open cluster of stars in the heart of the Orion Nebula contains several O-type stars of which the O6pe V star* θ^1 *Orionis C is the most massive. The optical light we see is the result from line emission of hydrogen (Balmer lines) and reflection of star light on dust grains.*

nature, and due to known elements such as nitrogen, neon, sulfer and argon[1]. The complexity of the spectrum of H II regions is nicely illustrated in Fig. 8.2, that shows the IR spectrum of K3-50A.

H II regions are formed when massive stars reach the main sequence and suddenly 'switch on' their UV radiation field. The gas and dust that remains from the star formation process is often still around and will be ionized by the UV radiation. Regions of ionized gas around O and early-B stars are therefore often called H II regions. In a sense, H II regions signal the end of star formation in a molecular cloud, as the photo-ionized gas is also heated, which causes over-pressurization through which the cloud is dispersed.

[1] See Table 2 of Lee & Hyung (2013) for a comprehensive overview of nebular lines in the H II region of the the planetary nebula NGC 6803).

Tielens, page 252

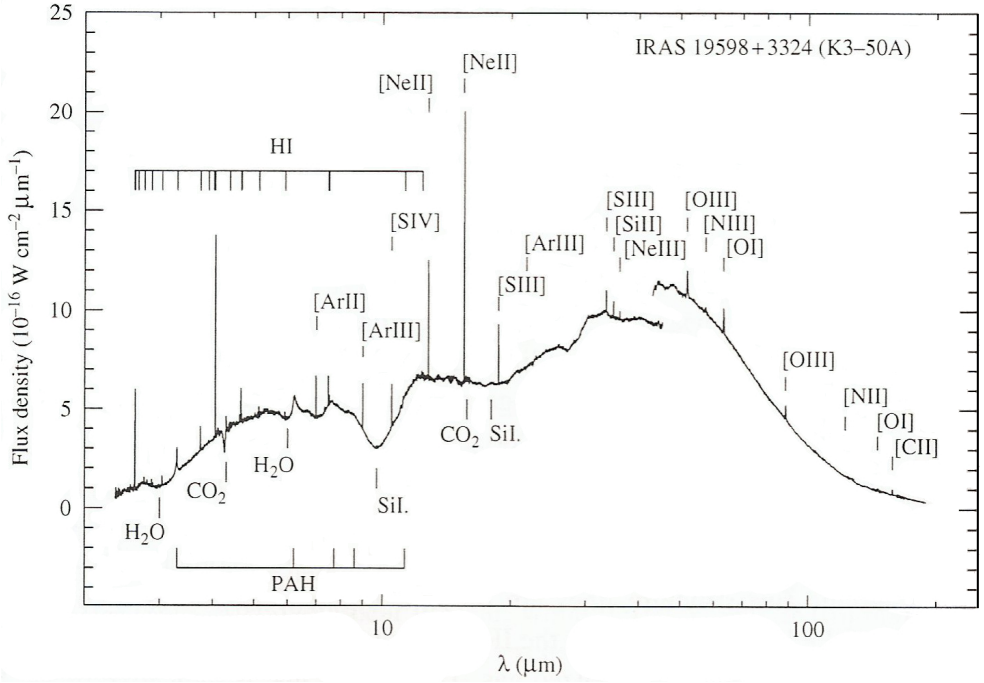


Figure 8.2: *Infrared spectrum of the dense (ultra-compact) H II gas-cloud surrounding the K3-50A high-mass star-forming region, measured by the Infrared Space Observatory. The distance to this source is ∼8.7 kpc. This spectrum shows a multitude of components. The continuum emission is due to warm dust. Superimposed, are IR emission bands due to PAH molecules. Absorption bands due to silicates and ices located in a foreground cold molecular cloud are also visible. The line spectrum shows hydrogen recombination lines and far-IR ionic fine structure lines originating in the H II regions, as well as those originating in the photo-dissociation region. The jump in the spectrum reflects the difference is aperture size between ISO's short wavelength and long wavelength spectrometer and is not intrinsic. Adapted from: Peeters et al. (2002), A&A 381, 571.*

Planetary nebulae have many things in common with H II regions: there too a gas/dust mixture (the stellar wind which was ejected when the star was still an asymptotic giant branch star) is ionized by the UV photons of the hot white dwarf. White dwarfs can reach much higher temperatures than main sequence stars, so in general one can expect higher ionization of the gas in planetary nebulae.

H II regions form a class of relatively well-studied objects that are usually classified as ultra-compact, compact, and extended H II regions (e.g. Habing & Israel 1979). Ultracompact H II regions (UCH IIs) have small sizes, of about 0.1 pc, and are located in the inner, high-pressure, parts of the parental molecular cloud. Compact H II regions have larger sizes, $0.1 - 0.3$ pc, and are carving their way out of the clouds. Extended H II regions, have sizes of up to several parsecs and they represent the mature state of these objects. Giant and supergiant

H II Class	Size [pc]	Density [cm^{-3}]	EM [pc cm^{-6}]	Ionized Mass [$M_\odot$]	Number of ionizing stars
Hypercompact	~ 0.003	$\gtrsim 10^6$	$\gtrsim 10^8$	$\sim 10^{-3}$	$\simeq 1$
Ultracompact	$\lesssim 0.05$	$\gtrsim 10^4$	$\gtrsim 10^7$	$\sim 10^{-2}$	~ 1
Compact	$\lesssim 0.5$	$\gtrsim 5 \times 10^3$	$\gtrsim 10^7$	~ 1	~ 1
Classical	~ 10	~ 100	$\sim 10^2$	$\sim 10^5$	few
Giant	$\sim 50 - 100$	~ 30	$\sim 5 \times 10^5$	$10^3 - 10^6$	$\sim 10^2$
Starburst nuclei	> 100	~ 10	$\sim 10^5$	$10^6 - 10^8$	$\sim 10^3$

Table 8.1: *Physical parameters of H II regions: size, mean density, emission measure, and ionized mass. From: Franco et al. (2000), Astrophysics and Space Science 272, 169; Murphy et al. (2010), MNRAS, 405, 1560, and Tielens (2005).*

H II regions are observed in external galaxies, but they represent a conglomerate of many individual H II regions that have already photo-ionized a large fraction of their parental giant molecular clouds. Recent observational results suggest the possibility of a new, even more compact type: Hypercompact (HCH II) regions. Physical properties of the classes of H II regions are summarized in Table 8.1.

In the next two paragraphs, we largely follow Draine (2011). If the dust to gas ratio in UCH II and HCH II regions is approximately 'normal', these regions are expected to be strongly affected by radiation pressure, and should exhibit a shell-like morphology if they are static. In these dense H II regions, we expect the dust to absorb a significant fraction of the ionizing radiation, as well as a substantial fraction of the recombination radiation, particularly Lyα. Consequently, the dust can be quite warm, and indeed, these regions stand out as sources that are bright at 24 μm or even 10 μm.

If the stellar source of ionizing radiation is stationary relative to the gas, then the timescale for expansion of a HCH II or UCH II region is very short – of the order of the Strömgren sphere divided by the $\sim$15 km s^{-1} sound speed (see Sect. 8.6) – and we would expect these objects to be relatively rare. The observed number of sources is, however, larger than expected. Several reasons may be responsible for this. First, it may be due to motion of the star relative to the gas: in the direction of motion of the star, the H II region ceases to expand when the expansion velocity of the gas is equal to the velocity of the star relative to the gas. In this scenario, the ionized gas should have a 'cometary' appearance: flattened on the leading edge of the H II region, with a 'tail' trailing behind the star. This morphology is sometimes seen (see Fig. 8.3). Second, some of the UCH II regions appear to be cases where a disk or other dense structure near the star is gradually being ablated by photo-ionization, providing a reservoir of gas to replace the gas removed by the expanding ionized outflow. If a disk is involved, the outflows may be bipolar. In case the dense structure is an ensemble of small dense neutral (molecular)

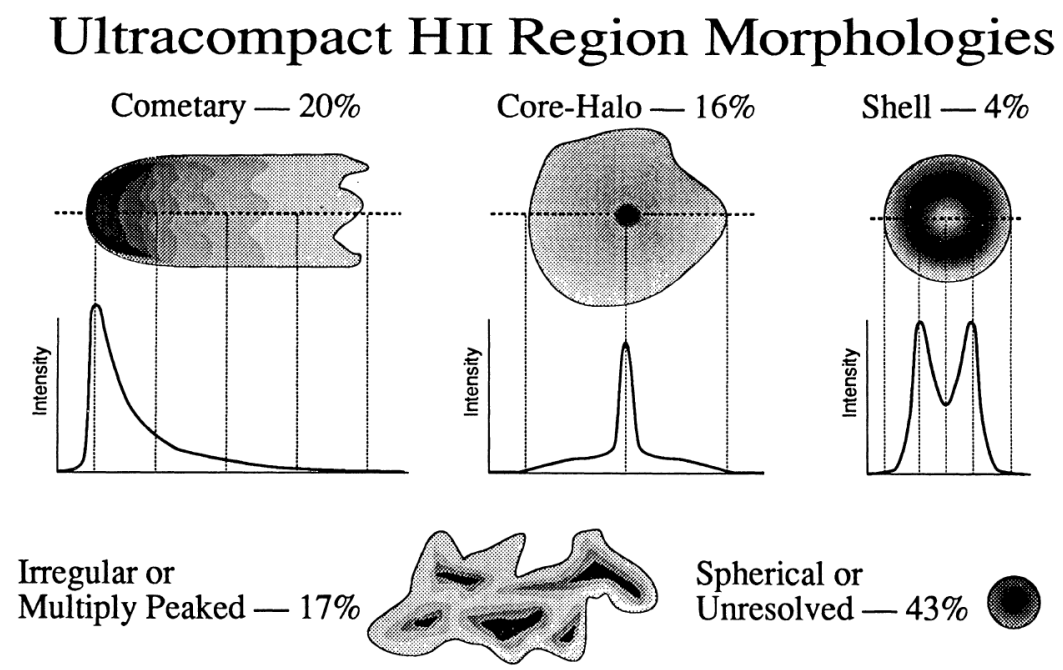


Figure 8.3: *Schematic pictures of basic UCH II region morphologies seen in high-resolution VLA observations. The spatial resolution was ∼0".4 and structures larger than ∼10" were not well imaged. Percentages give the relative number of sources in each class. Some of the spherical or unresolved sources may show different structures when observed at higher angular resolution, and a few non-cometary sources may actually be cometary sources viewed along their axis of symmetry. The appearance of a source may change with the wavelength of observation: the central cavities seen in the cometary and shell-shaped sources will not be seen at wavelengths where the gas is optically thick. From: Wood & Churchwell, 1989, ApJS, 69, 831.*

condensations, the stellar wind and UV flux of the young massive star may ablate and photoionize these 'blobs', thereby providing a continuous supply of ionized gas. (This is known as the mass-loaded stellar wind model). Third, newborn O or early-B stars are still embedded in remnants of their natal clouds. Their stellar winds will sweep up these remains in a shell and trap the ionizing radiation until the mass in the shell cannot absorb all of the ionizing photons emitted by the star (or other effects break up these shells) and ionizing radiation can stream out (Geen & de Koter 2022).

A nice review of UCH II regions is provided by Hoare et al. (2007). In these lecture notes, we only discuss spherical H II regions; for models of cometary and bipolar H II regions, see e.g. Redman et al. 1998, MNRAS, 298, 33.

8.2 Ionization and recombination rates

In Section 5.3 we considered transitions between two bound levels. Here, we will – again briefly – summarize transitions between a bound lower level l and a free upper level κ. The transition $l \rightarrow \kappa$ is equivalent to an ionization; the transition $\kappa \rightarrow l$ to a recombination recombination. The number of ionizations per second per cm^3 by radiative and collisional processes are

$$-\text{ Radiative ionizations} \quad : \quad \frac{\mathrm{d}n_l}{\mathrm{d}t} = n_l\, R_{l\kappa} = n_l\, 4\pi \int_{\nu_\circ}^{\infty} \alpha_{l\kappa}\, \frac{J_\nu}{h\nu}\, d\nu \tag{8.1}$$

$$-\text{ Collisions} \quad : \quad \frac{\mathrm{d}n_l}{\mathrm{d}t} = n_l\, C_{l\kappa} = n_l\, n_\mathrm{c}\, q_{l\kappa} \tag{8.2}$$

Three processes can result in the transition of the electron from the free state to the lower state. The number of such de-excitations (also in cm^{-3} s^{-1}) for spontaneous recombinations, stimulated recombinations and collisional recombinations are

$$-\text{ Spon. rec.} \quad : \quad \frac{\mathrm{d}n_\kappa}{\mathrm{d}t} = n_\kappa\, R_{\kappa l}^{\mathrm{spon}} = n_\kappa \left(\frac{n_l}{n_\kappa}\right)^* 4\pi \int_{\nu_\circ}^{\infty} \frac{\alpha_{l\kappa}}{h\nu} \frac{2h\nu^3}{c^2}\, e^{-h\nu/kT}\, d\nu \tag{8.3}$$

$$= n_\kappa\, n_\mathrm{c}\, \alpha_{i,\kappa-1}^{\mathrm{RR}}(T) \tag{8.4}$$

$$-\text{ Stim. rec.} \quad : \quad \frac{\mathrm{d}n_\kappa}{\mathrm{d}t} = n_u\, R_{\kappa l}^{\mathrm{stim}} = n_\kappa \left(\frac{n_l}{n_\kappa}\right)^* 4\pi \int_{\nu_\circ}^{\infty} \frac{\alpha_{l\kappa}}{h\nu} J_\nu\, e^{-h\nu/kT}\, d\nu \tag{8.5}$$

$$-\text{ Collisions} \quad : \quad \frac{\mathrm{d}n_\kappa}{\mathrm{d}t} = n_\kappa\, C_{\kappa l} \quad = n_u\, n_\mathrm{c}\, q_{\kappa l} \tag{8.6}$$

R and C are again the radiative and collisional rate in sec^{-1}. J_ν is the mean intensity of the radiations field, and $q_{l\kappa}$ and $q_{\kappa l}$ the cross-sections for collisions in cm^3 s^{-1}. The ratio $(n_l/n_\kappa)^*$ refers to the LTE equation of state Eq. 2.74. $\alpha_{l\kappa}$ is the extinction coefficient for photoionization or *photoionization cross-section* per particle in energy level i. It has dimension cm^2. The same remarks as to the nature of the collision particle 'c' as made in Section 5.3 apply here. Notice that one may express the number of spontaneous recombinations in terms of a recombination coefficient in cm^3 s^{-1}. In H II regions, where the collision partners are free electrons, $\alpha_{i,\kappa-1}^{\mathrm{RR}}(T)$ is a function that decreases with increasing temperature, as it is more difficult for an ion to capture a faster moving electron (see e.g. table 8.2). In ISM studies it is common to use this recombination coefficient formalism.

8.3 Ionization and recombination of hydrogenic atoms and ions

The ionization equilibrium of hydrogen is determined by the balance between photo-ionizations and recombinations of electrons with hydrogen ions. Let us assume that the nebula is composed of hydrogen only. What then is the ionization condition of H? Given the large excitation energies of hydrogen, it is safe to assume that all the neutral hydrogen is in the ground state, i.e. $n_0 = n^\circ$, where n° is the density of neutral hydrogen. The condition is then

$$n_0\, 4\pi \int_{\nu_\circ}^{\infty} \alpha_\nu^{\rm bf} \frac{J_\nu}{h\nu}\, d\nu = n^+ n_{\rm e}\, \alpha_{\rm A}(T), \tag{8.7}$$

where n^+ is the ionized hydrogen density, $n_{\rm e}$ the electron density, and T the kinetic temperature. The number of ionizations per second per cm^3 is shown on the left, and is the product of the ground state population density n_0, the *ionization cross-section* $\alpha_\nu^{\rm bf}$ in cm^2, and the number of ionizing photons $4\pi\, J_\nu/h\nu$ that illuminate the gas. The integration starts at $\nu_\circ$, the frequency of the ionization edge. For the ground state of hydrogen we have $\nu_\circ = 13.6\,$eV or 912 Å. Photons with higher energies are referred to as Lyman-continuum photons. For hydrogen and other ions with one electron (so-called hydrogenic ions)

$$\alpha_\nu^{\rm bf}(n, Z) = 7.91 \times 10^{-18} \frac{n}{Z^2}\, g_{\rm II}(\nu, n) \left(\frac{\nu_\circ}{\nu}\right)^3, \tag{8.8}$$

in which $\nu_\circ$ is the frequency of the ionization edge. Z is the nuclear charge of the atom or ion and $g_{\rm II}(\nu, n)$ is the bound-free Gaunt factor, which close to the ionization edge is ~ 1. The extinction is thus proportional to ν^{-3} for $\nu > \nu_\circ$. For more complicated ions, that have more valance electrons (for instance Fe I which has an electron shell that is half filled and that provides a multitude of valance electrons and valance cavities), the smooth decline is disrupted by resonances which produce all sorts of peaks in ionization cross-section that need to be determined in experiments.

On the right of Eq. (8.7) the number of recombinations per second per cm^{-3} are given. This is the product of the particles involved in the collision process, i.e. the ionized hydrogen density n^+ and the free electrons density. The proportionality constant is the *total recombination coefficient* of hydrogen to all bound levels $\alpha_{\rm A}$, in cm^3 s^{-1} (see Table 8.2). This function decreases with increasing temperature, as it is more difficult for an ion to capture a faster moving electron.

If $n = n^\circ + n^+$ is the hydrogen density, $q = n^\circ/n$ is the fraction of neutral hydrogen. It then follows that $n^+ = n_{\rm e} = (1 - q)\, n$. We further assume that the radiation field in the nebula is controlled by a star at a large distance. If we adopt the star to be an isotropically radiating sphere, it follows from Eq. (2.7) and (2.20) that

$$J_\nu(r) = \frac{1}{4\pi}\mathcal{F}_\nu(r) = \frac{1}{4\pi}\frac{L_\nu}{4\pi r^2}. \tag{8.9}$$

Table 8.2: *Recombination coefficients of hydrogen as a function of temperature in cm^3 s^{-1}: $\alpha_{\rm A} \equiv \sum_{i=1}^{\infty} \alpha_{i,{\rm H}}^{\rm RR}$ gives the total number of recombinations to all levels; $\alpha_{\rm B} \equiv \sum_{i=2}^{\infty} \alpha_{i,{\rm H}}^{\rm RR}$ gives the total number of recombinations to all levels above the ground level. From: Storey & Hummer, for $n_{\rm e} = 1000$. The fit formula is from Draine (2011), in which $T_4 = T/10\,000$ and Z the net charge of the ion. It provides good approximations in the range $30\,{\rm K} \leq T/Z^2 \leq 30\,000\,{\rm K}$.*

T	$\alpha_{\rm A}/\ 10^{-13}$	$\alpha_{\rm B}/\ 10^{-13}$
3 000	9.74	6.74
5 000	6.83	4.53
10 000	4.17	2.59
15 000	3.11	1.84
20 000	2.51	1.43
30 000	1.84	0.991
T	$4.13\,Z^2\,\left(T_4/Z^2\right)^{-0.7131-0.0115\,\ln\left(T_4/Z^2\right)}$	$2.54\,Z^2\,\left(T_4/Z^2\right)^{-0.8163-0.0208\,\ln\left(T_4/Z^2\right)}$

Writing for the total number of Lyman continuum photons leaving the star each second

$$Q_0 \equiv \int_{\nu_\circ}^{\infty} \frac{L_\nu}{h\nu}\, d\nu, \tag{8.10}$$

and approximating the extinction coefficient of photoionization with a constant (frequency independent) value $\overline{\alpha}_\nu^{\rm bf}$, we obtain for the fraction of neutral hydrogen

$$\frac{q}{(1-q)^2} = \frac{4\pi r^2 \alpha_{\rm A} n}{\overline{\alpha}_\nu^{\rm bf} Q_0}. \tag{8.11}$$

If the argument on the right-hand side is much smaller than unity, the left-hand side may be approximated by q. Let us consider a point in the nebula with a typical particle density $n = 10$ cm^{-3}, at a distance of 5 pc from an O7 V star having $\log Q_0 = 48.7$ (see Table 8.3). We then find – using $\overline{\alpha}_\nu^{\rm bf} \sim 6 \times 10^{-18}$ cm^2 (see Eq. 8.8) and $\alpha_{\rm A} \sim 4 \times 10^{-13}$ cm^3 s^{-1} (see Table 8.2) – for the fraction of neutral hydrogen $q \sim 4 \times 10^{-4}$. In other words: hydrogen is almost completely ionized.

How far can the Lyman-continuum photons typically travel before they are absorbed in a photo-ionization interaction? To this end, we calculate the mean free path at the distance 5 pc from the star. Using Eq. 2.45, we find

$$\Delta r \sim \frac{1}{\overline{\alpha}_\nu^{\rm bf}\, n^\circ} = \frac{1}{\overline{\alpha}_\nu^{\rm bf}\, q\, n} \sim 13.6\,{\rm pc}. \tag{8.12}$$

For this low particle density the medium is essentially optically thin for $\nu > \nu_\circ$. Be cautious that for $q \ll 1$, $\Delta r \propto 1/(rn)^2$, therefore for a higher density medium and further out from the central star this may no longer be the case.

Spectral type	$T_{\rm eff}$ [K]	$R_\star$ [$R_\odot$]	$\log L$ [$L_\odot$]	M [$M_\odot$]	v_∞ [$\mathrm{km\,s^{-1}}$]	$\log \dot{M}$ [$M_\odot \mathrm{yr}^{-1}$]	$\log Q_0$	$E_{\rm w}/E_{Q_0}$ %
O3	44900	13.8	5,84	58	3600	-5.6	49.6	1.2
O4	42900	12.4	5.67	47	3300	-5.7	49.4	1.3
O5	40900	11.2	5.49	38	3000	-6.0	49.2	0.8
O5.5	39900	10.6	5.41	34	2900	-6.1	49.1	0.8
O6	38700	10.1	5.32	31	2800	-6.2	49.0	0.7
O6.5	37900	9.6	5.23	28	2700	-6.3	48.9	0.7
O7	36900	9.2	5.14	25	2500	-6.4	48.7	0.7
O7.5	35900	8.7	5.05	23	2400	-6.5	48.6	0.7
O8	34900	8.3	4.96	21	2300	-6.6	48.4	0.8
O8.5	33900	7.9	4.86	19	2200	-6.7	48.3	0.7
O9	32900	7.5	4.77	17	2100	-6.8	48.1	0.8
O9.5	31900	7.2	4.68	16	2000	-6.9	47.9	0.9

Table 8.3: *Stellar parameters as a function of spectral types for luminosity class V stars (Martins et al. 2005, A&A 436, 1049; parameters for luminosity classes III (giants) and I (supergiants) are available as well). Terminal wind velocities, v_∞, and mass-loss rates, $\dot{M}$, are from Sternberg et al. (2000, ApJ 599, 1333). The final column provides the ratio between wind energy $E_{\rm w} = \dot{M} v_\infty^2/2$ and energy in ionizing photons E_{Q_0}, using the Lyman threshold frequency 3.3×10^{15} Hz (or 91,2 nm) to convert number of ionizing photons Q_0 to energy.*

8.4 Strömgren sphere

A source that is producing a finite number of ionizing photons can of course not ionize an infinitely large volume. Therefore, if the star is located in a sufficiently large gas cloud, there must be a boundary to the H II region. Somewhere there will be a zone in which hydrogen is recombining and outside of which there is an H I region. The thickness of the transition zone will be small because when hydrogen starts to recombine the optical depth in the Lyman continuum rapidly increases, effectively preventing ionizing radiation to penetrate deeper into the medium. Hence we may estimate the thickness of the recombination zone by equating it to the mean free path at the point where half of the hydrogen is recombined, i.e. at $q = 0.5$. We find (see Eq. 2.45) using the numbers we have used above

$$\Delta r \sim \frac{1}{\overline{\alpha}_\nu^{\rm bf}\, n^\circ} = \frac{1}{\overline{\alpha}_\nu^{\rm bf}\, q\, n} \sim 0.01\,\mathrm{pc}. \tag{8.13}$$

This is much smaller than the typical radius of an H II region (see below). H II regions thus have a sharp boundary, referred to as the *ionization front*. Inside of the boundary hydrogen is almost completely ionized; outside of the boundary hydrogen is almost completely neutral.

Strömgren sphere – optically thin Lyman continuum

How big is an H II region? Let us assume a homogeneous spherical cloud of hydrogen gas with a star in the center, and that the cloud is transparent for Lyman continuum photons. In very good approximation, the sphere up to the ionization front is almost fully ionized. We can thus write $n^+ \simeq n_\mathrm{e} \simeq n$. If R_S is the radius of the H II region, the balance between the number of ionizations per second and the number of recombinations per second is

$$Q_0 = \frac{4\pi}{3} R_\mathrm{S}^3 \, \alpha_\mathrm{A} \, n_\mathrm{e} \, n^+ = \frac{4\pi}{3} R_\mathrm{S}^3 \, \alpha_\mathrm{A} \, n^2. \tag{8.14}$$

Hence, the radius of the sphere is

$$R_\mathrm{S} = \left(\frac{3 Q_0}{4\pi \, \alpha_\mathrm{A} \, n^2} \right)^{1/3}. \tag{8.15}$$

This radius is referred to as the *Strömgren radius*, in honor of the pioneer of this field, and the volume encompassed by R_S as the *Strömgren sphere*. It shows that the higher the gas density, the smaller is the H II region.

If we fill in the numbers of our standard example (see above), we find $R_\mathrm{S} = 10.1$ pc. Indeed, the Strömgren radius is much larger than the thickness of the ionization front.

Strömgren sphere – optically thick Lyman continuum

In the previous discussion we assumed the Strömgren sphere to be optically thin and the radiation field in the nebula to be completely controlled by the central star. However, at Lyman continuum frequencies $\nu \geq \nu_\circ$ the nebula very likely is not optically thin. Therefore, we now allow the mean intensity to have a diffuse component as a result of Lyman continuum radiation emitted in the nebula. The local volume emission coefficient η_ν for radiation generated by recombinations to the ground state is given by

$$4\pi \int_{\nu_\circ}^{\infty} \frac{\eta_\nu}{h\nu} \, d\nu = n^+ \, n_\mathrm{e} \, \alpha_{1,\mathrm{H}}^{\mathrm{RR}}, \tag{8.16}$$

where $\alpha_{1,\mathrm{H}}^{\mathrm{RR}}$ is the recombination coefficient to the ground level of hydrogen.

Suppose the nebula is optically thick for Lyman radiation, such that no Lyman continuum photons can escape. Each and every Lyman photon produced by the diffuse radiation field will be absorbed elsewhere in the nebula. For the mean intensity of the diffuse component of the radiation field, J_ν^d, this implies

$$n_0 \, 4\pi \int_\mathrm{nebula} \alpha_\nu^\mathrm{bf} \, \frac{J_\nu^\mathrm{d}}{h\nu} \, dV = 4\pi \int_\mathrm{nebula} \frac{\eta_\nu}{h\nu} \, dV \tag{8.17}$$

where the integration is over the entire volume of the nebula. If the mean free path of the photons is small, that is, if the photons are absorbed close to where they are created, one may

even write

$$J_\nu^{\rm d} = \frac{\eta_\nu}{n_0\,\alpha_\nu^{\rm bf}}. \tag{8.18}$$

This is referred to as the *on-the-spot approximation*. For nebulae having particles densities that are not too small this is a reasonable assumption.

As a result of absorptions (read: photo-ionizations) Lyman continuum radiation that originates directly from the star gets weakened by a factor $\exp[-\tau_\nu(r)]$, where τ_ν is the radial optical depth

$$\tau_\nu(r) = \int_0^r n_0\,\alpha_\nu^{\rm bf}\,dr. \tag{8.19}$$

In a formal sense, the integration of course needs to start at $R_\star$ in stead of at the origin. However, as we will show below, things work out more elegantly if we accept this minute error. For the mean intensity of the stellar component of the radiation field we may write (using also Eq. 8.9)

$$J_\nu^{\rm s}(r) = \frac{1}{4\pi}\,\frac{L_\nu}{4\pi r^2}\,e^{-\tau_\nu}. \tag{8.20}$$

The total radiation field in the nebula is then the sum of the stellar and the diffuse component, i.e. $J_\nu = J_\nu^{\rm s} + J_\nu^{\rm d}$. Substituting Eq. (8.18) and (8.20) in ionization equilibrium (8.7) then yields

$$n_0 \int_{\nu_\circ}^\infty \alpha_\nu^{\rm bf}\,\frac{L_\nu}{h\nu}\,\frac{e^{-\tau_\nu}}{4\pi r^2}\,d\nu = n^+\,n_{\rm e}\left[\alpha_{\rm A} - \alpha_{\rm 1,H}^{\rm RR}\right] = n^+\,n_{\rm e}\,\alpha_{\rm B}, \tag{8.21}$$

where we have used Eq. (8.16). We have introduced the recombination coefficient $\alpha_{\rm B} = \sum_2^\infty \alpha_{i,{\rm H}}^{\rm RR}$ to represent the total number of recombinations to all levels above the ground level (see Table 8.2). The physical meaning of the above equation is that in an optically thick nebula ionizations caused by the stellar radiation field are in equilibrium with recombinations to excited levels of hydrogen. Recombinations to the ground level produce Lyman continuum photons that (in the on-the-spot approximation) are almost on-the-spot reabsorbed and thus have no effect on the ionization equilibrium.

To determine the radius of the ionized region we use that $d\tau_\nu = n_0\,\alpha_\nu^{\rm bf}\,dr$ (see Eq. 8.19) and integrate over distance. This yields

$$\int_{\nu_\circ}^\infty \frac{L_\nu}{h\nu}\left[\int_0^\infty e^{-\tau_\nu}\,d\tau_\nu\right]\,d\nu = \int_{\nu_\circ}^\infty \frac{L_\nu}{h\nu}\,d\nu = \int_0^\infty N^+\,n_{\rm e}\,\alpha_{\rm B}\,4\pi r^2\,dr \tag{8.22}$$

By letting the optical depth run from zero to infinity we have assured that all Lyman continuum photons are 'used up'. To find a simple expression for the size of the ionized region we again assume that the gas is fully ionized up to the Strömgren radius $R_{\rm S}$, and that the gas outside of this region is neutral. As we have seen, we expect a classical H II region to have a sharp boundary implying that our assumption is quite reasonable. Using Eq. (8.10) it follows that

$$Q_0 = \frac{4}{3}\pi R_{\rm S}^3\,\alpha_{\rm B}\,n^2 \tag{8.23}$$

The meaning of this result is that the total number of ionizing photons emitted by the star (per second) is exactly balanced by the total number of recombinations to excited states of all hydrogen atoms within a Strömgren sphere. For the Strömgren sphere we find

$$R_{\rm S} = \left(\frac{3Q_0}{4\pi\,\alpha_{\rm B}\,n^2}\right)^{1/3} = 1.6\,{\rm pc}\,\left(\frac{Q_0}{10^{50}\,{\rm s}^{-1}}\right)^{1/3}\,\left(\frac{n}{10^3\,{\rm cm}^3}\right)^{-2/3} \tag{8.24}$$

If we again fill in the numbers of our standard example (see above) and adopt $\alpha_{\rm B} = 2 \times 10^{-13}$ cm^3 s^{-1}, we find $R_{\rm S} = 12.7$ pc. The Strömgren sphere is now larger because only recombinations to exited levels matter. As there are less such events relative to recombinations to all levels, a larger volume can be kept ionized.

The two assumptions that we have made to estimate the size $R_{\rm S}$ of the H II region are the on-the-spot approximation and that the medium is either fully ionized (within the Strömgren sphere) or fully neutral (outside of the Strömgren sphere). In doing so we could avert having to solve the equation of transfer explicitly.

Ionization bounded and density bounded

Note that we have assumed that the ionizing star is in a sufficiently large gas cloud, that is, that the edge of the Strömgren sphere is determined by the region where hydrogen recombines. This is referred to as *ionization bounded*. It is however possible that the nebula contains insufficient hydrogen for all Lyman continuum photons to 'be used'. In that case we refer to the H II region as *density bounded*.

What is the total mass of ionized hydrogen gas in an ionization bounded nebula. This is

$$M = \frac{4\pi}{3}\,R_{\rm S}^3\,m_{\rm amu}\,n = \frac{Q_0\,m_{\rm amu}}{\alpha_{\rm B}\,n} \simeq 417\,M_\odot\,\left(\frac{Q}{10^{50}\,{\rm s}^{-1}}\right)\,\left(\frac{10^3\,{\rm cm}^{-3}}{n}\right), \tag{8.25}$$

where we have assumed the Lyman continuum to be optically thick, hence have adopted $\alpha_{\rm B}$. For our standard example we find $M = 2089\,M_\odot$.

Case A and Case B recombination

While we are in the process of defining things: in case the nebula is optically thin in all recombination lines, i.e. all radiation produced by recombination processes in the nebula is able to escape freely, one speaks of *Case A*. H II regions that fulfill the Case A requirement can only contain a relatively small amount of gas. So little gas, as a matter of fact, that they will be hard to observe.

Nebulae containing appreciable amounts of gas will rapidly develop large optical depths in the hydrogen Lyman lines. We can easily estimate this from the ratio of the extinction coefficients

per particle in the ground state for line- and continuum radiation. For the extinction coefficient of lines we have

$$\chi_\nu = \alpha_{lu}(\nu)\, n_l = \frac{h\nu}{4\pi}\, B_{lu}\, \phi_\nu\, n_l = \frac{\pi e^2}{m_\mathrm{e} c}\, f_{lu}\, \phi_\nu\, n_l, \tag{8.26}$$

where $(\pi e^2/m_\mathrm{e} c)$ = 0.02654 cm^2 hz and the dimensionless quantity f_{lu} is the *oscillator strength*. Comparison of the collisional cross section at line center $\alpha_{lu}(\nu_\circ)$, assuming a Doppler function for the profile function ϕ_ν, with the continuum cross section Eq.(8.8) at the ionization boundary yields, after substitution of constants

$$\frac{\tau(\mathrm{Ly}\, n)}{\tau(\mathrm{Ly{-}edge})} = 14.68\, \frac{\lambda_{1n}\, [\mathring{A}]\, f_{1n}}{g_\mathrm{II}(\nu_\circ, n)\, (T/10\,000)^{1/2}}, \tag{8.27}$$

where n is the principle quantum number of the upper level of the line that is considered. For a characteristic temperature $T = 10\,000$ K and $g_\mathrm{II} \sim 1$ we find that Lyα has about a 10^4 times larger optical depth than the continuum at the Lyman ionization boundary. For an ionization bounded nebula with τ(Ly-edge) ~ 1 one thus obtains τ(Lyα) $\sim 10^4$, τ(Lyβ) $\sim 10^3$, τ(Ly8) $\sim 10^2$, and τ(Ly18) ~ 10. For a typical nebula a better approximation than Case A therefore is the opposite assumption, i.e. that it is optically thick in all Lyman lines. This assumption is called *Case B*.

The above two limiting cases have been described in 1938 by Menzel and Baker and are therefore often called Menzel & Baker Case A and Case B. It should not be forgotten that the realistic situation in a nebula is likely in between these two limits. For fairly low lying transitions in the Lyman series Case B will hold, while for the higher transitions, i.e. $n \to \infty$ and small $\tau(\mathrm{Ly}\, n)$, the situation will be more similar to Case A.

8.5 Feedback of massive stars on the interstellar medium

Above, we showed that a massive hot star will photo-ionize the interstellar gas in its immediate surroundings. This particular process is only one way in which these stars affect their environment. Before exploring some of these so called *feedback* effects in more detail, we provide a brief overview of massive star feedback effects.

- The intense Lyman-continuum radiation field of hot massive stars creates ionized hydrogen bubbles, i.e. H II regions. H II regions in molecular clouds are bounded by photo-dissociation regions where molecular hydrogen is dissociated by Lyman-Werner band photons (see Sect. 10.3). For stars hotter than about 40 000 K the H II region will coincide with an He II region; the hottest of massive stars may also produced fully ionized helium bubbles, i.e. He III regions.

- In photo-ionizing the interstellar gas, hot massive stars heat the gas from about 10-50 K to about 10^4 K (see Sect. 9.1), destroying dust and (complex) molecules that may be present in the ambient medium.

- The increase in temperature (by a factor 200-1000) and the number of free particles (by a factor of 2) due to the photo-ionization process itself, greatly increase the gas pressure in the H II region. This ionized gas can not be contained and will expand. A shock front develops at the (moving) Strömgren radius, where interstellar matter accumulates in a thin shell (see Sect. 8.6). When the surface density of this shell material is high enough, it may fragment and cause new star formation.

- Mass loss from hot massive stars in super-sonic outflows too create a pressure on the ambient medium, causing an expanding shock wave and the accumulation of interstellar matter in a shell.

- Massive stars terminate their existence in core-collapse supernovae, extremely violent events. Gas is injected in the ISM at speeds of about 10 000 $\mathrm{km\,s^{-1}}$, clearing out channels all the way to the nearest edge of their host galaxies. In this way material may be lost from these galaxies altogether, and hydrogen ionizing photons may be streaming into the intergalactic medium through these channels. At early cosmic time, this process may have contributed strongly to the re-ionization of the IGM.

- Through stellar winds of hot and cool massive stars and their supernovae, the ISM is enriched in nuclear processed material, including the elements from which terrestrial planets and life are build. Part of this material enters the ISM in the form of dust. The ionizing radiation from hot massive stars may play a role in the chemistry that takes place in molecular clouds and in proto-planetary disks around nearby newly formed low-mass stars, potentially either aiding in creating complex chemistry or shutting such chemistry down.

8.6 Expansion of an H II region

In discussing the expansion of the H II region we closely follow Ward-Thompson & Whitworth (2011). We shall assume that, when a massive star forms, the output of hydrogen-ionizing photons Q_0 builds up instantaneously (i.e. the star 'switches on' abruptly) and then stays at a constant value during the main-sequence lifetime of the star. Also, we ignore the presence and impact of stellar winds (see Geen & de Koter 2021 for the effect of winds). The subsequent evolution can then be discussed in terms of three consecutive phases. We start by analyzing the initial static phase, then the intermediate phase, in which the H II region undergoes dynamical expansion, and lastly the final phase, in which the expansion halts. We note that some of the most massive stars may become supernova before the final phase is reached.

Initial static ionization phase

In the first phase, the initially neutral gas is static and the hydrogen-ionizing photons from the central star ionize it out to the Strömgren radius Eq. (8.24). The timescale on which the initial ionization equilibrium is established (i.e. the timescale for the hydrogen ionizing output from the central star to come into balance with the total rate of recombination in the H II region) is of order the recombination timescale – i.e. the average time that a free proton or electron has to wait before it finds a mate with which to recombine, given by

$$t_{\rm rec} = \frac{1}{\alpha_{\rm B}\, n} \simeq 160\,{\rm yr}\,\left(\frac{n}{10^3\,{\rm cm}^{-3}}\right)^{-1}. \tag{8.28}$$

We have assumed here that in the time $t_{\rm rec}$ the density of hydrogen in the H II region has not changed. This is only true if the timescale for expansion of the H II region driven by its overpressure relative to the surroundings is much longer. This dynamical expansion time may be estimated from

$$t_{\rm exp} \simeq \frac{R_{\rm S}}{a} \simeq 1.7\times10^5\,{\rm yrs}\,\left(\frac{Q_0}{10^{50}\,{\rm s}^{-1}}\right)^{1/3}\,\left(\frac{n}{10^3\,{\rm cm}^{-3}}\right)^{-2/3} \tag{8.29}$$

where

$$a = \left(\frac{kT}{m_{\rm amu}}\right)^{1/2} = 9.1\left(\frac{T}{10^4\,{\rm K}}\right)^{1/2}\quad {\rm km\,s}^{-1} \tag{8.30}$$

is the isothermal sound speed in the H II region, and we have assumed a temperature of 10^4 K.

For $t_{\rm rec} << t_{\rm exp}$ to hold it follows that

$$n \gg 10^{-6}\,{\rm cm}^{-3}\,\left(\frac{Q_0}{10^{50}\,{\rm s}^{-1}}\right)^{-1}. \tag{8.31}$$

This is clearly satisfied by the neutral gas in regions like giant molecular clouds where massive stars, and hence H II regions form. For a patch of gas in which a massive star forms of density $n = (1-5)\times10^4$ cm^{-3}, we find from Eq. (8.24) a Strömgren radius for our O7 V star emitting $\log Q_0 = 48.7$ of $R_{\rm S} = 0.14 - 0.04$ pc. This is typical for (ultra-)compact regions (see Table 8.1). These are often strong radio continuum sources because of the free-free emission the ionized gas produces, but they are invisible at shorter wavelengths because of the huge extinction by interstellar dust in the optical and near-IR.

The dynamical expansion phase

Even when all Lyman continuum photons are used to counteract recombinations the H II region will still expand. The densities in the H II region and its immediate surroundings are similar, but because of the higher temperature in the H II region the pressure can be hundreds of times higher. In order to reach pressure equilibrium with the surrounding medium the

H II region must therefore expand, which implies that the density will decrease (recall that $R_{\rm S} \propto n^{-2/3}$; see Eq. 8.24). This in turn reduces the number of recombinations per unit time (which is $\propto n^2$; see e.g. Eq. 8.7), so that more hydrogen gas can be ionized. Indeed the total number of atoms that can be ionized is $\propto n\, R_{\rm S}^3 \propto 1/n$.

We use Eq. (8.23) to write for the evolution of the number density of nucleons as a function of time

$$n(t) = n_0 \left(\frac{R}{R_0}\right)^{-3/2} = n_\circ \left(\frac{R}{R_0}\right)^{-3/2}, \tag{8.32}$$

where n_0 (do not get confused with the ground level H population) and R_0 reflect the situation at the end of the initial static ionization phase. For simplicity and clarity of notation, we refer to the time-dependent Strömgren radius as R (rather than $R_{\rm S}(t)$), and to the number density of the undisturbed medium in which the H II is expanding as $n_\circ$.

Since the expansion of the H II is initially at a speed $a \sim 10\,{\rm km\,s^{-1}}$, which is much greater than the sound speed $a_\circ \sim 0.3\,{\rm km\,s^{-1}}$ in the neutral gas outside the H II region, the ionization front at the edge of the H II region is preceded by a shock front, which sweeps the neutral gas up into a dense shell. We assume that the shock compression is very large (see below) such that the dense layer between the ionization front and the shock front is always very thin. Hence, both fronts have radial speed $\dot{R} = dR/dt$.

We assume that the pressure in the shocked gas is equal to the ram pressure of the undisturbed neutral gas flowing into the shock front at speed $\dot{R}(t)$

$$p_{\rm s} = \rho_\circ\, \dot{R}^2 = n_\circ\, m_{\rm amu}\, \dot{R}^2, \tag{8.33}$$

where $n_\circ$ is the number density and $\rho_\circ$ the density in the ambient gas[2]

We further assume that the pressure in the shocked gas is equal to the thermal pressure in the H II region

$$p_{\rm s} = \rho\, a^2 = n\, m_{\rm amu}\, a^2. \tag{8.34}$$

Finally, we assume that, although the pressures in the shocked layer and in the H II region are functions of time, they are uniform. This means that the sound travel time across the ionized region should be less than the timescale for a substantial change in the overal configuration. This may not be a good approximation in the early stages of expansion, where the ionized gas expands with a velocity comparable to a. It gets progressively better as the expansion progresses and must in any case be used in order to make a simple analysis. See Fig. 8.5 for a schematic representation of an H II region. Eliminating $p_{\rm s}$ between Eqs. (8.33) and (8.34), we obtain

$$n(t) = n_\circ\, \frac{\dot{R}^2}{a^2}. \tag{8.35}$$

[2]Notice the subtle difference in notation: n_0 is the number density in the H II region at the end of the initial static ionisation phase and $n_\circ$ the number density in the ambient medium.

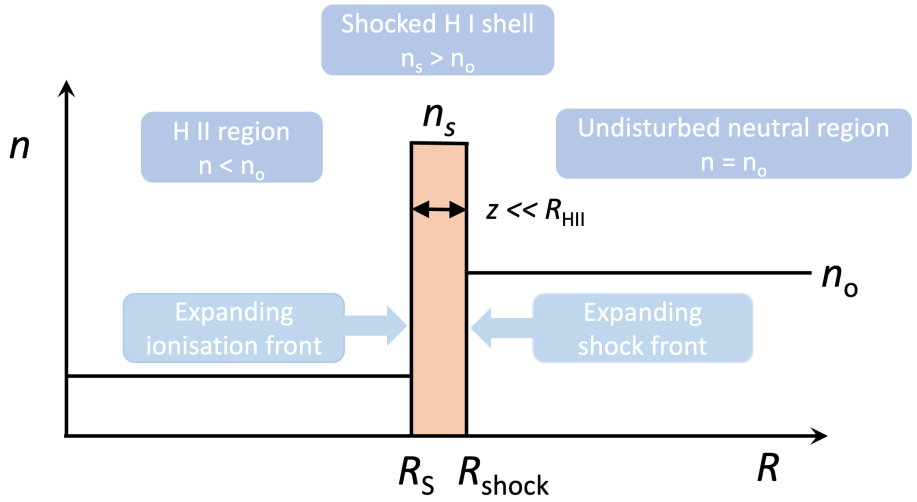


Figure 8.4: *Variation of density with radius when a shock wave driven by an expanding H II bubble moves into surrounding neutral material of constant density $n_\circ$ and creates a geometrically thin swept-up shell of neutral gas. The layer of shocked neutral gas is thin because of the compression across the shock. In effect the Strömgren radius $R_{\rm S} = R_{\rm shock}$. Based on: D. Ward-Thompson & A.P. Whitworth (2011).*

Notice that at $t = 0$, the start of the dynamical expansion, $\dot{R} = a$; the initial velocity of expansion is the sound speed in the H II region, as already pointed about above. Then eliminating $n(t)$ between Eqs. (8.32) and (8.35), we have

$$\dot{R}\, R^{3/4} = a\, R_0^{3/4}. \tag{8.36}$$

This can be integrated to give

$$R^{7/4}(t) = R_0^{3/4} \left(R_0 + \frac{7}{4} a t \right) \sim \frac{7}{4} R_0^{3/4}\, a\, t \tag{8.37}$$

The last expression in this equation gives the limiting behavior at late times. This is an adequate expression once $t \gg R_0/a \sim t_{\rm exp}$; where $t_{\rm exp}$ is short compared to the main-sequence lifetimes of OB stars. Therefore, through most of its lifetime the H II region's Strömgren radius is given approximately by

$$R(t) \simeq R_0^{3/7} \left(\frac{7}{4} a t \right)^{4/7}. \tag{8.38}$$

The asymptotic stage

However, the dynamical expansion phase can't go on forever. Eventually the expansion of the H II region will reduce its pressure to a value equal to the pressure in the undisturbed neutral

gas, and the expansion will then halt (provided the star lives long enough). If at this stage the density and the radius of the H II region have values $n_{\rm f}$ and $R_{\rm f}$, pressure balance with the undisturbed neutral gas requires

$$\rho_{\rm f}\, a^2 = \rho_\circ\, a_\circ^2, \qquad \text{such that} \qquad \frac{n_{\rm f}}{n_\circ} = \left(\frac{a_\circ}{a}\right)^2 = \frac{T_\circ}{T} \sim \frac{1}{1000}. \tag{8.39}$$

Here we have put the temperature of the neutral H I gas at $T_\circ = 10\,{\rm K}$ and of the ionized H II gas at $T = 10^4\,{\rm K}$ [3]

So, the radius $R_{\rm f}$, for which the expansion will halt is

$$R_{\rm f} = \left(\frac{n_{\rm f}}{n_\circ}\right)^{-2/3} R_0 = \left(\frac{T}{T_\circ}\right)^{2/3} R_0 \sim 100\, R_0, \tag{8.40}$$

where R_0 is the initial Strömgren radius Eq. (8.24). Using Eq. 8.38, the time to reach $R_{\rm f}$ is

$$t_{\rm f} = \frac{4}{7}\left(\frac{R_0}{a}\right)\left(\frac{R_{\rm f}}{R_0}\right)^{7/4} = \frac{4}{7}\, t_{\rm exp}\left(\frac{R_{\rm f}}{R_0}\right)^{7/4} = 23\,{\rm Myr}\left(\frac{Q}{10^{50}\,{\rm s}^{-1}}\right)^{1/3}\left(\frac{n}{10^3\,{\rm cm}^3}\right)^{-2/3}, \tag{8.41}$$

where in the second equality we have substituted expression Eq. 8.29. Note that $t_{\rm exp}$ is not a physically relevant timescale in this problem (its just convenient to use the expression) as expansion of the shockfront starts from R_0 after the static ionization phase has ended. The timescale $t_{\rm f}$ is comparable to the main-sequence lifetime of OB stars. We conclude that the most massive O stars, which have the highest Q_0 and shortest lifetimes, end their lives whilst the H II region is still over-pressured and expanding. The least massive OB stars, with relatively low Q_0 and relatively long lifetimes will reach the asymptotic stage of H II region expansion. When the source of ionizing photons is gone, it will take about $t_{\rm rec}$ yrs (see Eq. 8.28) for the gas to become neutral again.

The swept-up neutral gas at the boundary of an H II region

As well as the ionized gas we should also consider the shell of neutral gas swept up between the ionization front and the shock front. For a mature H II region, the mass of the swept-up shell of neutral gas at the edge is usually much greater than the mass of the ionized gas.

The mass in the H II region is given by

$$M_{\rm H\,II}(R) = \frac{4}{3}\pi R^3\, m_{\rm amu}\, n = \frac{4}{3}\pi R_0^3\, m_{\rm amu}\, n_0 \left(\frac{R}{R_0}\right)^{3/2} = M_0\left(\frac{R}{R_0}\right)^{3/2}, \tag{8.42}$$

[3] In our analysis of H II region expansion we have not considered the role of the mean molecular weight μ. The pressure $p = nkT = \rho kT/\mu m_{\rm amu} = \rho a^2$, is proportional to the number of free particles n, or, alternatively phrased, to $1/\mu$. If the H II gas consists of hydrogen only, $\mu = 0.5$, and if the ambient medium consists of molecular hydrogen, $\mu = 2$. This impacts pressure considerations throughout the discussion of H II region expansion. For the temperature ratio in Eq. (8.39) this would imply a ratio of 1/4000 rather than 1/1000.

where in the second equality we have used Eq. (8.32) and M_0 is the mass of the H II region at the end of the initial static ionization phase (see also Eq. 8.25). It follows that if the Strömgren spheres moves out a distance dR, the mass that is added to the H II region is

$$dM_{\rm H\,II} = \frac{3}{2}\,M_0\,\left(\frac{R}{R_0}\right)^{1/2}\frac{1}{R_0}\,dR = 2\pi R_0^{3/2}\,m_{\rm amu}\,n_0\,R^{1/2}\,dR, \tag{8.43}$$

where we (as discussed above) assume that the density and temperature – hence the pressure – of this freshly ionized material are quickly homogenized with that of the rest of the ionized gas. Over the same interval dR, the shockfront just downstream of the ionization front (which we place at the same radius as the ionization front) encounters an amount of interstellar material

$$dM_{\rm ISM} = 4\pi R^2\,m_{\rm amu}\,n_\circ\,dR, \tag{8.44}$$

implying

$$\frac{dM_{\rm HII}}{dM_{\rm ISM}} = \frac{1}{2}\left(\frac{R_0}{R}\right)^{3/2}. \tag{8.45}$$

So, at the start of the dynamical expansion phase half of the mass that is swept up at the interface with the ambient medium is added to the H II region and half to that of the thin shell of shocked gas. At later times hardly any mass enters the H II region and almost all is added to the shocked shell. For instance, at $R = 14\,R_0$ less than 1% is added to the H II region and over 99% is added to the thin shocked shell. The mass in the shocked shell is

$$M_{\rm shell} = \frac{4}{3}\pi R^3\,m_{\rm amu}\,n_\circ - \frac{4}{3}\pi R^3 m_{\rm amu}\,n = \frac{4}{3}\pi R^3 m_{\rm amu}\,n_\circ\left[1-\left(\frac{R_0}{R}\right)^{3/2}\right]. \tag{8.46}$$

where we have used Eq. (8.32). Indeed, at later times it is safe to approximate the mass in the shell as being all of the mass that has been swept up.

As the surface-density of the shell grows it becomes increasingly gravitationally unstable, and eventually it should break up into collapsing fragments. These collapsing fragments may produce a new generation of stars, and this is usually referred to as propagating, or triggered, star formation. If some of the new stars are massive, and therefore excite new H II regions, the process can repeat itself, and we speak of self-propagating star formation.

8.7 The impact of the stellar wind on the interstellar gas

All throughout their evolution massive stars loose mass in a spherical stellar wind. The rate of mass loss for the most massive stars in their main sequence phase is in the range $10^{-7} - 3\times10^{-6}\ M_\odot{\rm yr}^{-1}$. The material in the wind is accelerated to a maximum outflow speed or terminal velocity of $2000 - 3600\ {\rm km\,s^{-1}}$ (see Table 8.3).

8.8 Gas density diagnostics: the emission measure

Let us assume that the fraction of all recombinations of hydrogen that lead through the transition $u \rightarrow l$ is p_{ul}. Then, the volume emission coefficient for the line $u \rightarrow l$ is (see also Eq. 8.7)

$$\eta_\nu = \frac{h\nu}{4\pi}\, p_{ul}\, \alpha_{\rm A}\, n^+\, n_{\rm e}\, \phi_\nu = \eta_\circ\, \phi_\nu \tag{8.47}$$

The expression is the multiplication of the number of de-excitations from $u \rightarrow l$ per cm^3 per second, per unit solid angle – which is why we need to devide by the total solid angle $\Omega = 4\pi$ – multiplied by the energy of the photon that is emitted $h\nu$, where $\nu = \nu_{lu}$. To get the emitted energy per hz, we multiply by the profile function for spontaneous emission ϕ_ν (which has dimension hz^{-1}).

Though to determine the temperature dependent p_{ul} requires a general approach, it is found that $p_{{\rm H}\alpha} \sim 0.3$, i.e. only about 30 percent of hydrogen recombinations produce an Hα photon. Similarly, $p_{{\rm H}\beta} \sim 0.1$, so only about 10 percent of hydrogen recombinations produce an Hβ photon.

If we observe an H II region along a line-of-sight where there is negligible background specific intensity, and if the emission from the H II region is optically thin, then (see Sect. 2.3)

$$I_\nu^{\rm obs} \simeq \int S_\nu(\tau_\nu)\, d\tau_\nu = \int \eta_\nu(s)\, ds, \tag{8.48}$$

where $I_\nu^{\rm obs}$ is the observed specific intensity and s is the spatial coordinate along the line-of-sight. It follows that the specific intensity integrated over the line proflle is

$$I^{\rm obs} = \int I_\nu^{\rm obs}\, d\nu = \int \left[\int \eta_\circ\, ds\right] \phi_\nu\, d\nu = \int \eta_\circ\, ds = \frac{h\nu}{4\pi}\, p_{ul}\, \alpha_{\rm A} \int n^+\, n_{\rm e}\, ds, \tag{8.49}$$

assuming the emitting medium has uniform temperature T (recall $\alpha_{\rm A} = \alpha_{\rm A}(T)$). The integral in the last equation is called the *emission measure* of the line-of-sight

$$\mathcal{EM}_{\rm H} = \int n^+(s)\, n_{\rm e}(s)\, ds. \tag{8.50}$$

If the emitting medium has uniform density, then $\mathcal{EM}_{\rm H} = n^+\, n_{\rm e}\, D$, where D is the length of the intercept which the line-of-sight makes with the emitting medium.

Thus a measurement of the total line specific intensity Eq. (8.49) enables us to estimate the emission measure of the emitting medium. Suppose, we also have an independent estimate of the linear size D of the emitting region, say from its distance d and its angular size α, via $D = d\,\alpha$. This assumes spherical symmetry, which is often a reasonable approximation. Then we can combine this with the emission measure to obtain $\langle n^+\, n_{\rm e}\rangle = \mathcal{EM}_{\rm H}/D$. Typically in an H II region we have $n^+ \sim n_{\rm e}$. Therefore

$$\langle n_{\rm e}^2\rangle = \langle (n^+)^2\rangle = \frac{\mathcal{EM}_{\rm H}}{D}. \tag{8.51}$$

So, from the study of hydrogen recombination lines, we can estimate the density of gas in an H II region and the total mass of ionized gas

$$M_{\mathrm{H\,II}} = \frac{4\pi}{3}\left(\frac{D}{2}\right)^3 m_{\mathrm{amu}}\, n^+, \tag{8.52}$$

which essentially is the total mass, as in the H II region $n^+ = n$.

Exercise 8.1

Consider a pure hydrogen gas and a situation in which only radiative and collisional ionizations R and C are important. The recombination of the medium is described by a recombination coefficient α.

a) Derive a quadratic equation describing the equilibrium ionized fraction $q = N^+/N$, where $N = N_0 + N^+$ is the sum of particles in the neutral state and in the ionized state.

b) Derive an approximate expression for the low-density limit $R \gg C \cdot N$.

c) Derive approximate expressions for the high-density limit $R \ll C \cdot N$, for the situations in which collisional ionizations are either (i) important or (ii) irrelevant.

Exercise 8.2

This could be a nice exam question. According to Martins et al. (2005, A&A 436, 1049) a 60 $M_\odot$ star produces $Q = 10^{49.7}$ ionizing photons per second early on in its life. Brott et al. (2011, A&A 530, 116) predict a total lifetime for such a star of about 4 Myr. We assume that Q remains constant throughout the life of the star. The medium in which the massive stars forms has density $N = 3 \times 10^4$ cm^{-3}. The sound speed a in the ionized part of the medium is $10\,\mathrm{km\,s^{-1}}$. In the neutral medium $a = 0.3$ kms.

a) Compute the Strömgren sphere radius at the end of the initial static ionization phase in pc.

b) Compute the Strömgren sphere radius at the end of the life of the star in pc.

c) How much mass has been swept up in the dense shell that has accumulated at the edge of the H II region once the star dies?

d) Linear stability analysis results in a characteristic length scale

$$\ell = \frac{2a^2}{G\Sigma} \tag{8.53}$$

for which the medium is most susceptible for instability. Σ is the surface density of the gas shell. Compute this length scale ℓ at the end of the star's life.

e) What would be the typical mass of a shell fragment that collapses at $t = 4\,$Myr? If collapse of part of the shell would occur earlier in the life of the star, would the mass of the fragment be larger / the same / or smaller than what you have just computed?

Exercise 8.3

This could be a nice exam question. Consider a spherical cloud, initially completely neutral, consisting of hydrogen with uniform number density $n^\circ = n$. At time $t = 0$, a hot star at the centre of the cloud starts emitting Lyman continuum photons at a constant rate Q_0. These cause a progressively larger sphere around the star to become ionized. The gas inside the *ionization front* at radius R will be mostly ionized, and at larger distances will be mostly neutral. Neglecting the width of this transition region, the position and speed of the front is found from

$$4\pi\, R^2\, n^\circ\, \frac{\mathrm{d}R}{\mathrm{d}t} + \frac{4}{3}\pi\, R^3\, \alpha_\mathrm{B}\, n^+\, n_\mathrm{e} = Q_0, \tag{8.54}$$

where n_e is the electron density, n^+ is the density of ions, and α_B is the recombination coefficient. The above equation is referred to as the *jump condition* and expresses that an ionizing photon either ionizes a neutral atom for the first time (the first left-hand side term) or compensates for a recombination within the ionized region (the second left-hand side term). We assume the gas inside the ionized region to be fully ionized, i.e. $n^+ = n_\mathrm{e} = n$, and beyond the ionization front, in the neutral zone, to be completely neutral, i.e. $n^\circ = n$. Then, Eq. 8.54 can be solved exactly.

a) Show that a radius evolution of the type

$$R(t) = a\,[1 - \exp(-bt)]^{1/3}\,, \tag{8.55}$$

is a valid solution and determine a and b. Link these constants to physical quantities relevant to this problem and discussed in this chapter.

b) Introducing the initial static ionization phase in section 8.6, it was stated that the timescale on which the initial ionization equilibrium was established is of the order of the recombination timescale t_rec. Given the results of a) and b), is this indeed a reasonable timescale?

c) Derive an expression for the speed of the ionization front.

Exercise 8.4

This could be a nice exam question. The Red Supergiant star Betelgeuse in the constellation of Orion travels through its local interstellar environment at a speed of $v_\star$ = 25 $\mathrm{km\,s^{-1}}$, creating a bow shock. The apex of this arc-like structure has a radius of 5.75 arcmin from the star which is at 200 pc. We assume the bow shock is observed edge on.

a) Compare the apparent apex radius in arcmin to the apparent radius (also in arcmin) of the moon.

b) Compute the apex radius in pc.

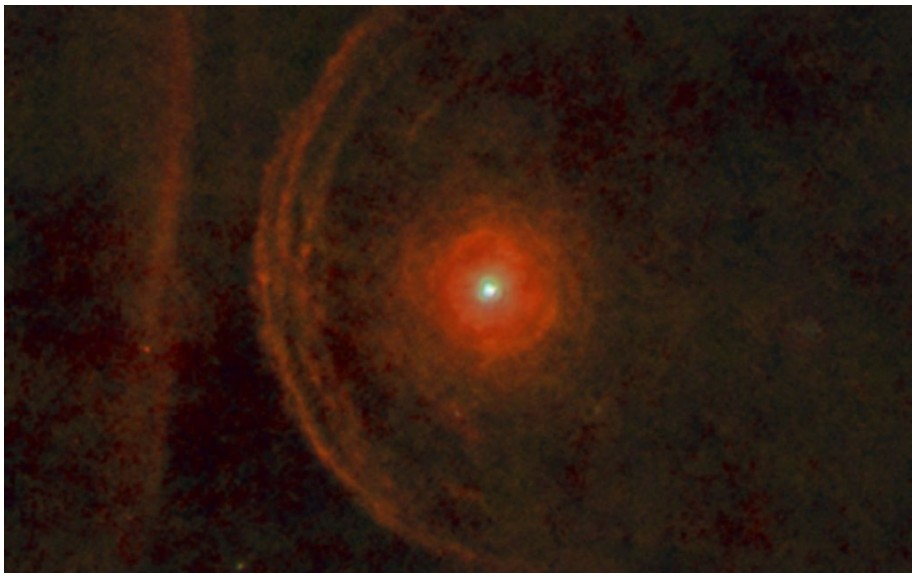


Figure 8.5: *Composite color image of the Herschel PACS 70, 100, 160 μm images of the Red Supergiant Betelgeuse (α Orionis) and its surroundings. Betelgeuse travels through its local intersteller environment at a speed of 25 $\mathrm{km\,s^{-1}}$, plowing up material forming an asymmetric arc or bow shock. The apex of the arc is at 5.75' from the star. Notice to the left a dust filament, thought to be unassociated with the bow shock. From: ESA-Herschel-PACS / Decin et al.*

For a bow shock, the *standoff distance* $R_{\rm so}$ is where the ram pressure of the stellar wind of the moving star balances that of the surrounding material. $R_{\rm so}$ is measured in the direction of motion of the star. The mass-loss rate is given by

$$\dot{M} = 4\pi\, r^2\, \rho_{\rm w}(r)\, v_{\rm w}, \tag{8.56}$$

where r is radial distance, $\rho_{\rm w}$ the wind density and $v_{\rm w}$ the wind velocity. For Betelgeuse $v_{\rm w} = 17\,\mathrm{km\,s^{-1}}$. We assume the ambient medium to be a neutral gas, for which the mean molecular weight is $\mu = 1.27$, and adopt a canonical value for the particle density $n_\circ = 1\,\mathrm{cm^{-3}}$.

c) Derive a formula for the mass-loss rate of Betelgeuse as a function of $R_{\rm so}$ and other relevant quantities.

d) Calculate the mass-loss of Betelgeuse in solar masses per year. The literature value is $\dot{M} = (3 \pm 1) \times 10^{-6} M_\odot \mathrm{yr}^{-1}$. How well does your result compare to this value? Which assumption(s) may contribute to the cause of the difference.

9

H II regions - part II: thermal state

We discuss the energetics of H II regions. We first review the heating and cooling processes of interstellar gas in these clouds. Then we discuss how one obtains key properties of H II regions, i.e. its temperature and density.

9.1 An overview of heating and cooling

What are the physical processes that add kinetic energy to, i.e. *heat*, the interstellar gas, and what processes extract kinetic energy from, i.e. *cool*, this gas? Together, these processes determine the thermal state of the gas.

In general terms, we can imagine two categories of heating processes in the ISM: microscopic scale processes (i.e. the absorption of photon energy by gas and dust) and large scale processes (mechanical heating, e.g. by SNe explosions and cloud-cloud collisions). While there are a number of energy sources with comparable energy densities in the ISM (see Sect. 1.3), the processes that couple the gas to the radiation field generally dominate. Below, we provide a brief overview of the most important of the heating and cooling processes, with a reference to the section where we provide more information. For more details see e.g. Tielens (2005) and Draine (2011).

Heating processes

- Photo-ionization: Here we need to distinguish between H II regions, where very energetic photons are available and H I regions, where no photons more energetic than 13.6 eV – the ionization potential of hydrogen – are present. In H II regions H ionization dominates the heating of the gas, while in H I regions photo-ionization of large (stable) molecules and small dust grains take over. When solids are photo-ionized, it is custom to speak of *photo-electric heating* or *photo-electric effect*.
- Cosmic rays: Of all cosmic rays, those having energies of 1-10 MeV are most efficient

in heating the gas. They can do so in two ways: (1) through interaction with bound electrons resulting in ejection of an energetic electron, that in turn may cause other (secondary) ionizations; and (2) transfer of kinetic energy to free electrons by elastic scattering. The latter process is only important in H II regions.

- ☐ Shockwaves: Shockwaves are common phenomena in the ISM, and occur whenever material moves at velocities exceeding the sound speed in the surrounding medium and the upstream material can not dynamically respond to the upcoming material until it arrives. The shock will then compress, heat, and accelerate the medium. Supernovae or cloud-cloud collisions are examples of events that may lead to shock heating.

Cooling processes

- ☐ Collisionally excited line radiation: The fundamental cooling process of interstellar gas is the collisional excitation of an atom, ion or molecule, followed by a radiative de-excitation. The collision partners can be electrons, H atoms or H_2 molecules, whichever are the most abundant. The process dominates the cooling in both H I and H II regions. Because the densities are low, forbidden transitions contribute – and even play a dominant role – in the cooling.
- ☐ Recombination radiation: Every time an electron recombines with an ion, the plasma loses the kinetic energy of the recombining electron. In the case of Case B recombination – which reflects the usual situation – only recombinations to levels above the ground level will produce photons that can escape the nebula. The process contributes (somewhat) to the cooling of H II regions.
- ☐ Free-free emission: Free electrons scattering off ions produce free-free emission. The photons created in this way typically have long wavelengths and may escape freely. As it requires an ionized plasma, this cooling mechanism need only be considered in H II regions.

9.1.1 Heating by photo-ionization

We examine the energy input by photo-ionizations by again considering a pure hydrogen nebula. At any specific point in the nebula, the heating rate in erg cm^{-3} s^{-1} is

$$G(\mathrm{H}) = n_0\, 4\pi \int_{\nu_\circ}^{\infty} \alpha_\nu^{\mathrm{bf}} \frac{J_\nu}{h\nu}\, (h\nu - h\nu_\circ)\, d\nu \tag{9.1}$$

$$= \alpha_{\mathrm{A}}\, n_{\mathrm{e}}\, N^+ \frac{4\pi \int_{\nu_\circ}^{\infty} \alpha_\nu^{\mathrm{bf}}\, (J_\nu/h\nu)\, (h\nu - h\nu_\circ)\, d\nu}{4\pi \int_{\nu_\circ}^{\infty} \alpha_\nu^{\mathrm{bf}}\, (J_\nu/h\nu)\, d\nu} \tag{9.2}$$

$$= \alpha_{\mathrm{A}}\, n_{\mathrm{e}}\, N^+\, E_{\mathrm{init}}, \tag{9.3}$$

where G is short for Gain (the letter Γ is used as well). The term $(h\nu - h\nu_\circ)$ reflects the thermal energy of the electron that is liberated from the hydrogen atom. It is this energy

times the number of ionizations per second per cm^3 that specifies the gain in thermal energy of the medium. In writing down the second equality, we have assumed that the nebula is in ionization equilibrium (see Eq. 8.7). In the third equality we have introduced the initial mean energy of the newly created photo-electron or *mean initial photon energy*, which depends on the shape of the ionizing radiation field, but not on the absolute strength of this field. This can be made explicit when we describe the stellar Lyman continuum radiation field with a diluted blackbody spectrum at a radiation temperature $T_{\rm rad}$ (see also Eqs. 2.7 and 5.45)

$$J_\nu = W(r)\, B_\nu(T_{\rm rad}). \tag{9.4}$$

Note that the dilution factor does not feature in $E_{\rm init}$, i.e. the absolute strength of J_ν does not matter. It is convenient to define the dimensionless ratio parameter ψ, such that (see also Eq. 2.63)

$$E_{\rm init} = \psi\, \frac{3}{2} kT_{\rm rad}. \tag{9.5}$$

We refer to ψ at the stellar surface as $\psi_\circ$, i.e. the case where no radiation has been absorbed yet. Though near the star the stellar radiation may already be attenuated (i.e. $W(r) \ll 1$), the shape will still be $J_\nu \propto B_\nu(T_{\rm rad})$, i.e. it is not yet distorted by frequency dependent absorption (that is $\propto \exp(-\tau_\nu)$).

For the important case of photo-ionization of hydrogen, we know from the discussion of the Strömgren sphere that – by definition – all of the photons with $h\nu > h\nu_\circ$ will produce a photo-ionization somewhere in the H II region. Since every stellar Lyman continuum photon is eventually absorbed in the H II gas, we need not weigh with the effect of the ionization cross section $\alpha_\nu^{\rm bf}$ and may set it to unity for all frequencies. We also need not care about the diffuse radiation field because the total energy that is gained through absorption of diffuse photons is exactly equal to the energy that is lost elsewhere in the H II region due to the recombinations that created these diffuse photons: a zero-sum game. This implies that the average of the photo-electron energy over the entire nebula is given by

$$\langle\psi\rangle\, \frac{3}{2} kT_{\rm rad} = \frac{\int_0^\infty (B_\nu/h\nu)\,(h\nu - h\nu_\circ)\; d\nu}{\int_0^\infty (B_\nu/h\nu)\; d\nu} \tag{9.6}$$

Values for the two cases $\psi_\circ$ and $\langle\psi\rangle$ are given in Table 9.1 for selected values of $T_{\rm rad}$. The important thing to note is that both $\psi_\circ$ and $\langle\psi\rangle$ are of order unity over a broad range of $T_{\rm rad}$.

A photo-ionization that is followed by a recombination to the ground-level implies that first radiation energy is converted into thermal energy, after which the gained thermal energy is converted back into radiation energy. If the Lyman continuum radiation field is optically thick, we may apply the on-the-spot approximation (see Sect. 8.4) stating that the recombinations occur 'on the spot', i.e. immediately after the ionization has occurred. If so, it is as if a photon scattering has occurred, i.e. no net energy is exchanged between the radiation field and the thermal pool of gas – again a zero-sum game. These combinations of events do not contribute to the energy balance and may be excluded, i.e. only photo-ionizations followed by recombinations to excited levels contribute to the heating. We may trivially exclude these

$T_{\rm rad}\ (K)$	4000	8000	16000	32000	64000
ψ_0	0.651	0.639	0.615	0.576	0.517
$\langle\psi\rangle$	0.701	0.734	0.799	0.920	1.103

Table 9.1: *Dimensionless factor providing the mean initial photon electron energy $E_{\rm init} = \psi\, kT_{\rm rad}$. The values of ψ_0 are lower than the corresponding values of $\langle\psi\rangle$ because of the ν^{-3} dependence of the photo-ionization cross-section, which means that photons near the ionization edge are absorbed close to the star; hence the radiation field hardens with distance from the star. Adapted from: Spitzer (1998); see also Tielens (2005).*

'scatterings' by replacing $\alpha_{\rm A}$ by $\alpha_{\rm B}$ in Eq. (9.3). The heating rate becomes

$$G({\rm H}) = \alpha_{\rm B}\, n_{\rm e}\, N^+\, \psi\, \frac{3}{2} kT_{\rm rad}. \tag{9.7}$$

9.1.2 Cooling by collisionally excited line radiation

Cooling of ionized gas is mainly achieved by means of collisional excitations followed by radiative de-excitations. Important transitions in this context are forbidden fine-structure transitions in O III, S III, O II, Ne II, and N II. For the energy level diagram of O III and N II see Fig. 3.2; that of O III is also displayed in Fig. 9.3. Notice that the forbidden transitions among the ^{3}P, ^{3}D and ^{1}S levels of O III separate out in two groups: (1) Those involving fine-structure levels of the ground state that are separated by less than ∼0.1 eV, and (2) transitions involving levels separated by 1−3 eV. As

$$E\ [{\rm eV}] = 11604.5\, T\ [{\rm K}]\,, \tag{9.8}$$

the first group involves levels that are separated by energies that correspond to 1000 K or less. The populations of these levels are therefore not very sensitive to the temperature of the gas when $T \sim 10^4$ K. Transitions involving levels of the second group are sensitive to the electron temperature and hence their excitation acts as a thermostat regulating the temperature of the ionized gas to $\sim 10^4$ K. In other words: because typically the fine-structure transitions do not deliver enough cooling, the gas temperature has to rise sufficiently (to $\sim 10^4$ K) to excite these higher levels.

Note that the cooling efficiency of these transitions depends on their critical density $n_{\rm cr}$, see Eq. (5.31) and (5.32).

The energy $L({\rm ion})$ that is released by this cooling in erg cm^{-3} s^{-1} is

$$L({\rm ion}) = h\nu_{lu}\, n_u\, A_{ul} = h\nu_{lu}\, n_{\rm e}\, (n_l\, q_{lu} - n_u\, q_{ul})\,, \tag{9.9}$$

where L is short for Loss (the letter Λ is used as well). l and u are the indices of the lower an upper level of the ion that is considered. The equation simply stems from the statistical

equilibrium equation $n_l\, n_{\rm e}\, q_{lu} \;=\; n_u\, n_{\rm e}\, q_{ul} + n_u\, A_{ul}$ and the emitted energy per cm^3 per second, $h\nu_{lu}\, n_u\, A_{ul}$, where $h\nu_{lu} = E_u - E_l$.

Obviously, the number of excitations to higher energies increases rapidly with increasing temperature, because the average velocity of the electrons increases with temperature. This means that the cooling rate also increases rapidly as temperature rises. A balance between heating and cooling follows from

$$G(H) \;=\; \sum_{\rm ions} L({\rm ion}), \tag{9.10}$$

where the sum is over all ions contributing to the cooling, and all relevant lines of these ions. In H II regions this balance is usually reached at about 10000 K. See Sect. 9.1.5 for a more precise description of thermal balance.

9.1.3 Cooling by recombination

What if cooling of our pure hydrogen nebula would occur by recombinations only? On average, the kinetic energy of the recombining electrons is $3kT_{\rm e}/2$ and the total loss of energy per unit volume per unit time is

$$L({\rm H}) = \alpha_{\rm B}\, n_{\rm e}\, N^+\, \frac{3}{2}\, kT_{\rm e}. \tag{9.11}$$

If only these recombinations cool the medium, thermal balance is given by $G(H) = L(H)$, hence

$$T_{\rm e} = \psi\, T_{\rm rad}. \tag{9.12}$$

Table 9.1 shows that $0.7 \lesssim \psi \lesssim 1.1$. For O-type stars $T_{\rm rad} \simeq 30 - 50\,$kK, yielding electron temperatures $T_{\rm e}$ in the nebula that are about the same (see Table 9.1). These are much higher values than the $\sim$10 kK that are usually observed in the H II regions. So, indirectly, this shows the importance of additional cooling processes – notably the cooling by collisionally excited line radiation.

9.1.4 Cooling by free-free emission

Hydrogen can cool through free-free transitions, where a free electron is decelerated in the electric field of a proton. Starting from Eq. (6.2), we may derive the cooling rate in erg cm^{-3} s^{-1} integrated over frequency

$$L({\rm ff}) = 4\pi \int_0^\infty \eta_\nu^{\rm ff}\, d\nu = 1.425 \times 10^{-27}\, Z^2\, T^{1/2}\, n_{\rm e}\, N^+, \tag{9.13}$$

where $Z = 1$ for ionized hydrogen. Cooling by free-free radiation is about equally important as cooling by recombination, i.e. in the present day universe it is not very important. In the early universe – when the chemical composition was H and He only – these two processes dominated the cooling in the ionized medium.

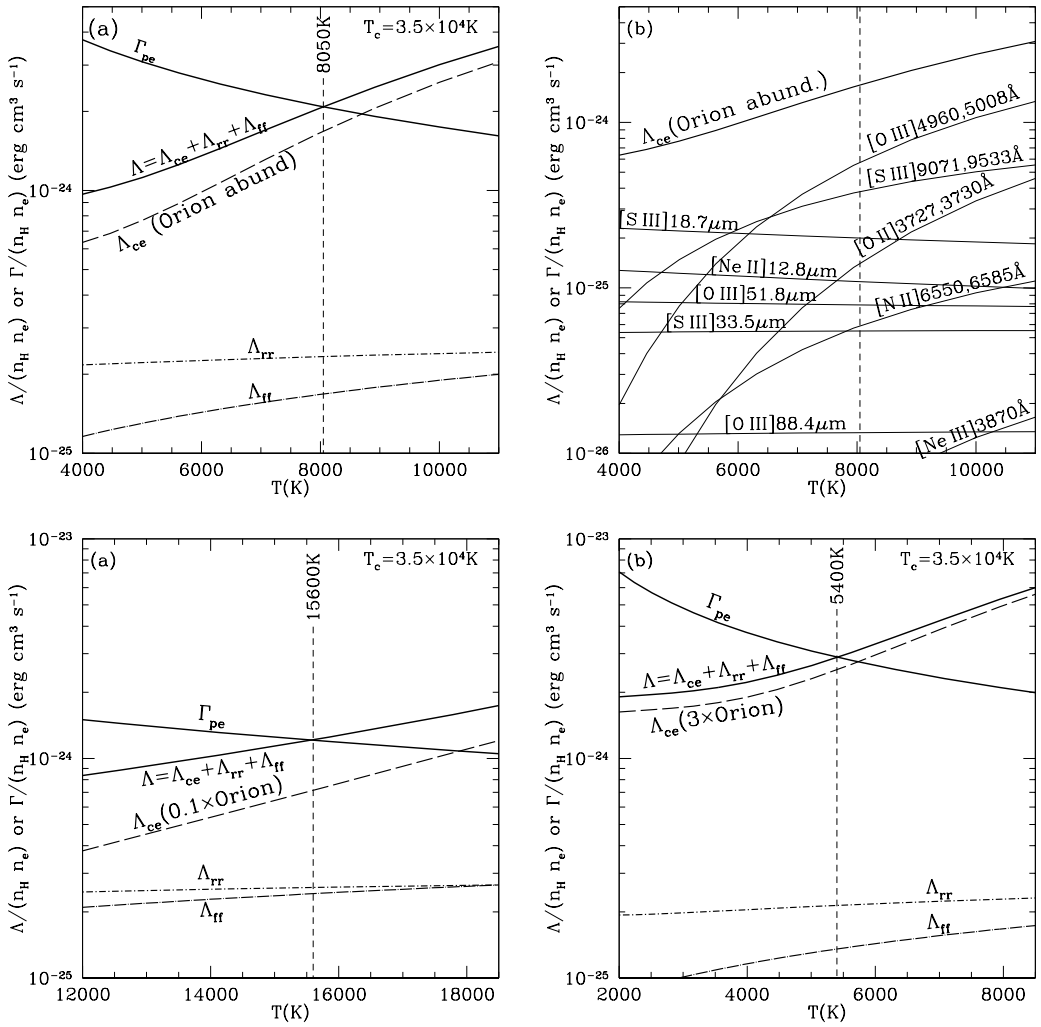


Figure 9.1: *Top (a): photo-electric heating function Γ_{pe} and radiative cooling function Γ as function of gas temperature T in an H II region with Orion-like abundances and density $n_{\rm H} = 4000$ cm^{-3}. Heating and cooling balance at $T \sim 8050$ K. Top (b): Contributions of collisionally excited individual lines to $\Lambda_{\rm ce}$. Bottom (a): same as in top left panel, but now for metal abundances that are 10% of solar. Bottom (b): same as in top left panel, but now for an enhanced metal abundance by a factor of three.*

9.1.5 Thermal balance between heating and cooling in H II regions

The gas will stabilize at a temperate where heating balances the total cooling, i.e. where

$$G(\mathrm{H}) \quad = \quad L(\mathrm{total}) = \sum_{\mathrm{ions}} L(\mathrm{ion}) + L(\mathrm{H}) + L(\mathrm{ff}). \tag{9.14}$$

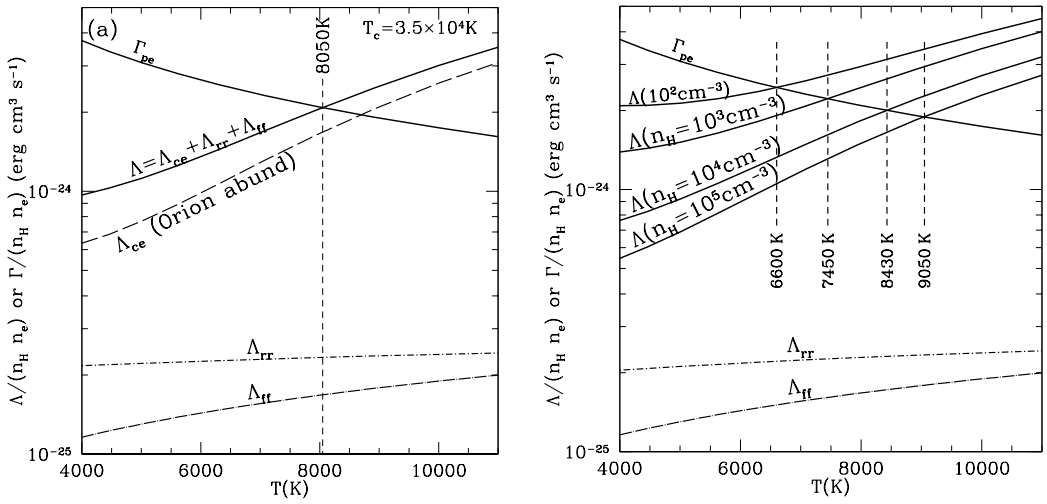


Figure 9.2: *Left: same as the top left panel of Fig. 9.1, for reference. Right: same as in left panel, but now for a range of hydrogen particle densities. As the gas density is increased from* 10^2 *cm*$^{-3}$ *to* 10^5 *cm*$^{-3}$*, the equilibrium temperature increases from 6600 K to 9050 K, because more and more of the cooling transitions become useless once the density gets above the critical densities of these forbidden lines.*

The top panels of figure 9.1 show the temperature dependence of $G(\mathrm{H})$ (in the figure Γ_{pe}), $L(\mathrm{ion})$ (in the figure Λ_{ce}), $L(\mathrm{H})$ (in the figure Λ_{rr}), $L(\mathrm{ff})$ (in the figure Λ_{ff}), and $L(\mathrm{total})$ (in the figure Λ) for gas with a solar composition, a density $n_{\mathrm{H}} = 4000$ cm^{-3}, and ionization similar to that in the Orion Nebula, assuming heating by a $T_{\mathrm{rad}} = 35000$ K blackbody. We find an equilibrium temperature of $T \sim 8050$ K.

The equilibrium temperature is sensitive to the abundance of the coolant species. Lowering the abundance will cause the equilibrium temperature to increase. This is shown in the lower panels of Fig. 9.1 – if we reduce the metal abundance by a factor of 10, as might be appropriate for a low-metallicity galaxy, the equilibrium temperature is raised to $\sim$ 15600 K. Conversely, raising the metal abundance by a factor of 3, as might be appropriate in the central regions of mature spiral galaxies, causes a drop to $\sim$ 5400 K.

For given gaseous abundances, the H II region temperature will also be sensitive to the gas density. When the density exceeds the critical density of some of the cooling levels, the cooling will be suppressed, and the equilibrium temperature will rise. For our test calculation it is found that if the density is increased from $n_{\mathrm{H}} = 10^2$ cm^{-3} to $n_{\mathrm{H}} = 10^5$ cm^{-3}, the thermal equilibrium temperature shifts from 6600 to 9050 K (see right panel Fig. 9.2).

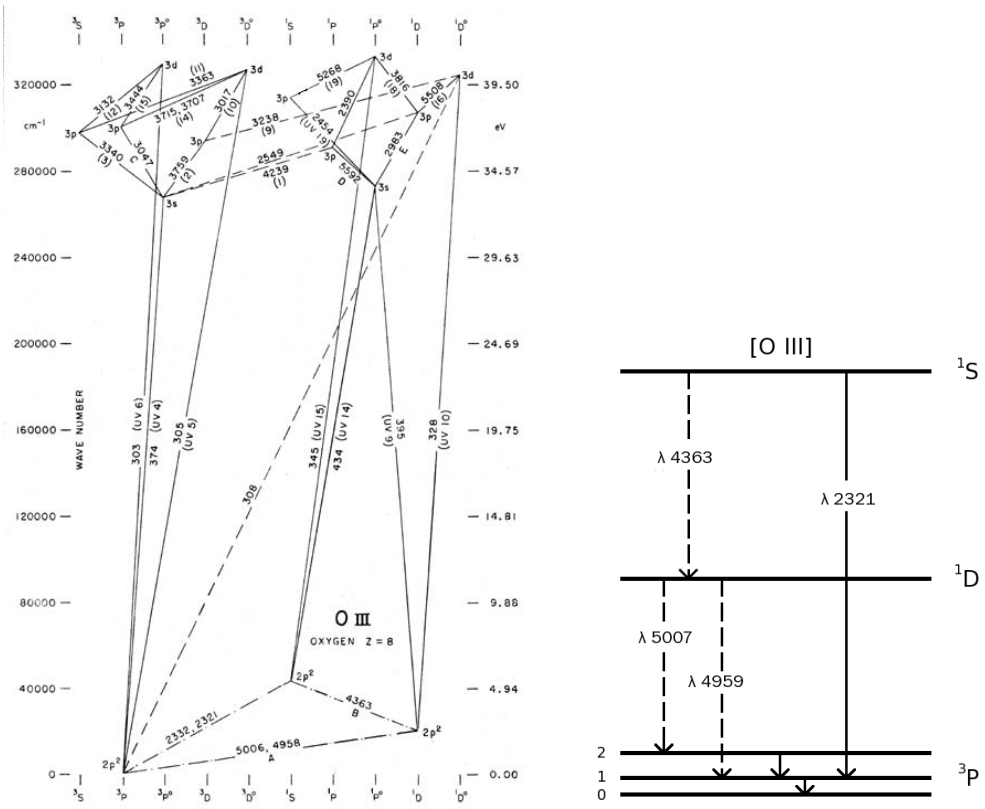


Figure 9.3: *Left: Energy level diagram of* O III*. Right: detail of the lowest levels of* O III *and the forbidden oxygen transitions (denoted by dashed lines) that are prominently visible in the emission spectra of nebulae. The* ^{3}P *term of the* $2p^2$ *level has three fine structure levels, i.e.* $^3\mathrm{P}_{0,1,2}$*. The relative energy differences between these fine structure levels have been strongly exaggerated for clarification purposes.*

9.2 T_e diagnostics: collisionally excited forbidden fine-structure lines

The forbidden emission lines of O III at λ4363, 4959 and 5007 Å (see Fig. 9.3) turn out to be very suited for constraining the temperature of the nebular gas. We assume that the electron density of the gas is much less than the critical density (see Eq. 5.31), such that collisional de-excitations are negligible.

To keep a clear view of the situation we refer to the ground state $^3\mathrm{P}_{0,1,2}$ as level 1, the first excited level $^1\mathrm{D}_2$ as level 2, and the second excited level $^1\mathrm{S}_0$ as level 3. This notation ignores that there are two[1] possible transitions from level 2 to 1, namely $^1\mathrm{D}_2 - {}^3\mathrm{P}_1$ (λ5007 Å) and

[1] See for the ΔJ selection rules Sec. 3.3. There is a third possibility, namely $^1\mathrm{D}_2 - {}^3\mathrm{P}_0$ (λ4931 Å), that can only occur by means of a quadrupole transition, but that is so weak that it can be ignored altogether.

$^1{\rm D}_2 - {}^3{\rm P}_2$ ($\lambda 4959$ Å), however, we will repair that later. The statistical equilibrium equation for level 2 is

$$\begin{aligned} n_2\,A_{21} &= n_1\,n_{\rm e}\,q_{12} + n_3\,A_{32} \\ &\simeq n_1\,n_{\rm e}\,q_{12}. \end{aligned} \tag{9.15}$$

This reflects that the n_2 level can be populated either by collisional excitations from the ground level ($n_1\,n_{\rm e}\,q_{12}$) or by spontaneous radiative de-excitations from level 3 ($n_3\,A_{32}$). Note that the second process is negligible relative to the first and can be safely ignored, simplifying the derivation. For level 3 it follows that

$$n_3\,(A_{31} + A_{32}) = n_1\,n_{\rm e}\,q_{13}. \tag{9.16}$$

This level can be populated by collisional excitations from the ground level ($n_1\,n_{\rm e}\,q_{13}$) or be de-populated by cascade to level 2 ($n_3\,A_{32}$) or level 1 ($n_3\,A_{31}$). Note that in this case we do not ignore the process $n_3 A_{32}$ as here it contributes significantly to the depopulation of level 3. The ratio of the populations of levels 2 and 3 is now given by

$$\frac{n_2}{n_3} = \frac{(A_{31}+A_{32})}{A_{21}}\frac{q_{12}}{q_{13}} = \frac{(A_{31}+A_{32})}{A_{21}}\frac{\Omega_{12}}{\Omega_{13}}\frac{e^{-E_{12}/kT}}{e^{-E_{13}/kT}} = \frac{(A_{31}+A_{32})}{A_{21}}\frac{\Omega_{12}}{\Omega_{13}}e^{+E_{23}/kT}, \tag{9.17}$$

where we have expressed the cross-sections for collisions q_{lu} using the *Maxwellian averaged collisional strength* Ω. In this description,

$$q_{lu}(T) = \frac{8.629\times10^{-6}}{g_l\,T^{1/2}}\,e^{-E_{lu}/kT}\,\Omega_{lu}, \tag{9.18}$$

and, as $n_l^{\rm LTE}\,q_{lu} = n_u^{\rm LTE}\,q_{ul}$,

$$q_{ul}(T) = \frac{g_l}{g_u}e^{+E_{lu}/kT}\,q_{lu}(T) = \frac{8.629\times10^{-6}}{g_u\,T^{1/2}}\,\Omega_{lu}. \tag{9.19}$$

Beware that $\Omega_{lu} = \Omega_{ul}$. Values of Ω are typically of order unity and for some important transitions are summarized in Table 9.2.

If we assume that all the lines of interest are optically thin, the observed line profile integrated flux is given by

$$\mathcal{F}_{lu} = \int_{\rm line}\left[\mathcal{F}_\nu(d) - \mathcal{F}_\nu{}^{\rm cont}(d)\right]\,d\nu \simeq \int_{\rm line}\mathcal{F}_\nu(d)\,d\nu \tag{9.20}$$

$$= \frac{1}{4\pi d^2}\int_V 4\pi\,\eta_{lu}\,dV = \frac{1}{4\pi d^2}\int_V h\nu_{lu}\,n_u\,A_{ul}\,dV, \tag{9.21}$$

$$\tag{9.22}$$

where $\mathcal{F}_\nu(d)$ is the observed flux from a nebula at distance d. The observed continuum flux $\mathcal{F}_\nu{}^{\rm cont}(d)$ is usually negligible for strong nebular lines (see Fig. 9.4). The line profile integrated emission coefficient η_{lu} (see also Eq. 5.37) multiplied by the total solid angle ($\Omega = 4\pi$)

Ion	Transition	λ_{lu} Å (or μm)	$\Delta E_{ul}/k$ K	A_{ul} s^{-1}	Ω_{ul} cm^{-3}	$n_{\rm cr}$ cm^{-3}
O II	$^2D_{5/2} - ^4S_{3/2}$	3728.8	39000	3.8×10^{-5}	0.82	1.3×10^3
	$^2P_{1/2} - ^2D_{5/2}$	7318.8	20000	2.0×10^{-2}	0.30	4.3×10^6
	$^2P_{1/2} - ^2D_{3/2}$	7329.6	20000	1.0×10^{-1}	0.28	4.3×10^6
	$^2P_{1/2} - ^4S_{3/2}$	2470.2	58000	2.3×10^{-2}	0.14	4.3×10^6
	$^2P_{3/2} - ^2D_{5/2}$	7319.9	20000	1.2×10^{-1}	0.74	6.3×10^6
	$^2P_{3/2} - ^2D_{3/2}$	7330.7	20000	6.1×10^{-2}	0.41	6.3×10^6
	$^2P_{3/2} - ^2S_{3/2}$	2470.3	58000	5.6×10^{-2}	0.28	6.3×10^6
N II	$^3P_1 - ^3P_0$	204μm	70	2.1×10^{-6}	0.41	4.5×10^1
	$^3P_2 - ^3P_1$	122μm	120	7.5×10^{-6}	1.12	2.8×10^2
	$^1D_2 - ^3P_2$	6583.4	22000	3.0×10^{-3}	1.47	8.6×10^4
	$^1D_2 - ^3P_1$	6548.1	22000	1.0×10^{-3}	0.88	8.6×10^4
	$^1S_0 - ^1D_2$	5754.6	25000	1.0	0.83	1.1×10^7
	$^1S_0 - ^3P_1$	3062.8	47000	3.3×10^{-2}	0.10	1.1×10^7
S II	$^2D_{3/2} - ^4S_{3/2}$	6730.8	21000	8.8×10^{-4}	2.76	3.6×10^3
	$^2D_{5/2} - ^4S_{3/2}$	6716.4	21000	2.6×10^{-4}	4.14	1.3×10^3
	$^2P_{1/2} - ^2D_{5/2}$	10370.5	14000	7.8×10^{-2}	2.20	9.8×10^5
	$^2P_{1/2} - ^2D_{3/2}$	10336.4	14000	1.6×10^{-1}	1.79	9.8×10^5
	$^2P_{1/2} - ^4S_{3/2}$	4076.4	35000	9.1×10^{-2}	1.17	9.8×10^5
	$^2P_{3/2} - ^2D_{5/2}$	10320.5	14000	1.8×10^{-1}	4.99	5.7×10^6
	$^2P_{3/2} - ^2D_{3/2}$	10286.7	14000	1.3×10^{-1}	3.00	5.7×10^6
	$^2P_{3/2} - ^4S_{3/2}$	4068.6	35000	2.2×10^{-1}	2.35	5.7×10^6
O III	$^3P_1 - ^3P_0$	88.4μm	160	2.7×10^{-5}	0.54	5.0×10^2
	$^3P_2 - ^3P_1$	51.8μm	280	9.8×10^{-5}	1.29	3.4×10^3
	$^1D_2 - ^3P_2$	5006.9	29000	2.0×10^{-2}	1.27	6.9×10^5
	$^1D_2 - ^3P_1$	4958.9	29000	7.0×10^{-3}	0.76	6.9×10^5
	$^1S_0 - ^1D_2$	4363.2	33000	1.7	0.58	2.4×10^7
	$^1S_0 - ^3P_1$	2321.4	62000	2.3×10^{-1}	0.10	2.4×10^7
S III	$^3P_1 - ^3P_0$	33.5μm	430	4.7×10^{-4}	3.98	1.4×10^3
	$^3P_2 - ^3P_1$	18.7μm	770	2.1×10^{-3}	7.87	1.2×10^4
	$^1D_2 - ^3P_2$	9530.9	15000	5.5×10^{-2}	3.86	6.2×10^5
	$^1D_2 - ^3P_1$	9068.9	16000	2.1×10^{-2}	2.32	6.2×10^5
	$^1S_0 - ^1D_2$	6312.1	23000	2.3	1.38	1.4×10^7
	$^1S_0 - ^3P_1$	3721.7	39000	8.4×10^{-1}	0.39	1.4×10^7

Table 9.2: *Optical and infrared cooling lines of ionized gas. All listed transitions are forbidden. Compiled by: Tielens (2005). Beware that* $\Omega_{ul} = \Omega_{lu}$.

is the total energy emitted by the spectral line per cm^{-3} per second. It thus needs to be integrated over the entire volume V of the nebular gas. If we assume a constant temperature, the *ratio of the line strengths* 2→1 and 3→2 is the same in each cubic centimeter of the nebula and we need not integrate over the full volume to obtain this ratio. We get

$$\begin{aligned}
\frac{\eta_{4959} + \eta_{5007}}{\eta_{4363}} &= \frac{h\nu_{4959}\, A_{4959} + h\nu_{5007}\, A_{5007}}{h\nu_{4363}\, A_{4363}} \frac{n_2}{n_3} \\
&= \frac{A_{2321} + A_{4363}}{A_{4959} + A_{5007}} \frac{\nu_{4959}\, A_{4959} + \nu_{5007}\, A_{5007}}{\nu_{4363}\, A_{4363}} \frac{\Omega_{12}}{\Omega_{13}}\, e^{+E_{23}/kT} \\
&= \frac{A_{2321} + A_{4363}}{A_{4363}} \frac{\overline{\nu}}{\nu_{4363}} \frac{\Omega_{12}}{\Omega_{13}}\, e^{+E_{23}/kT},
\end{aligned} \tag{9.23}$$

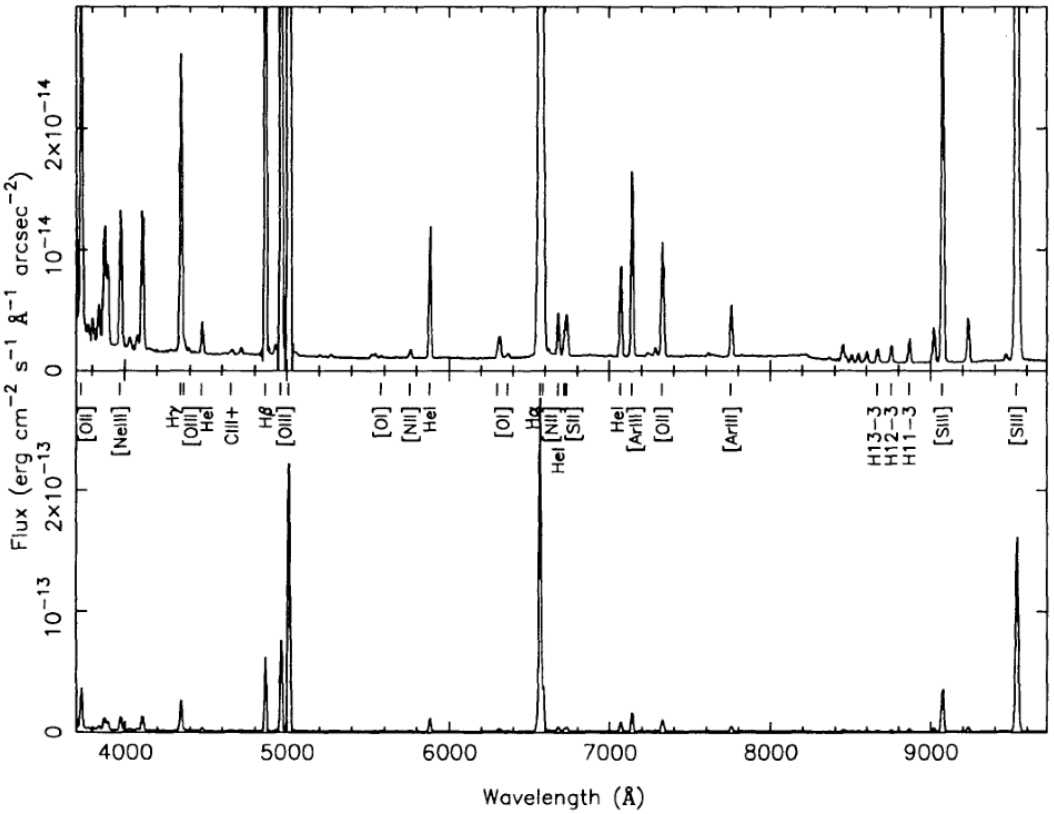


Figure 9.4: *The visible spectrum of a resolved part of the Orion Nebula* H II *region. In the upper panel, the spectrum has been scaled up to bring out the underlying continuum, which is weak relative to the spectral lines. From: Baldwin et al. (1991), ApJ 374, 580.*

where $\overline{\nu}$ is the Einstein A averaged frequency of the 2→1 transition. Notice that we now have split up the 2→1 transition in the lines $^1{\rm D}_2 - {}^3{\rm P}_1$ (λ5007 Å) and $^1{\rm D}_2 - {}^3{\rm P}_2$ (λ4959 Å).

This diagnostic is useful for temperatures ranging from $5\,000 - 20\,000$ K (or $\sim 0.5 - 2$ eV; as $1{\rm eV} = [k\,11605]$ K) and earns this sensitivity to the fact that the energy levels 2 and 3 are some distance apart. If indeed the distance between these levels would have been small, $\exp(E_{23}/kT) \sim 1$ and consequently the line flux ratios would have been insensitive to temperature. There are other ions that have a similar favorable positioning of energy levels and that can be used in a similar fashion to constrain the temperature of warm gaseous nebulae, for instance [N II], [O I], [Ne III], [S III]. Figure 9.5 shows the sensitivity of these diagnostics in the limit of low electron densities.

Equation (9.23) is a good approximation up to $n_{\rm e} \sim 10^5\,{\rm cm}^{-3}$. At higher electron densities collisional de-excitations start to play a role (see also Sect. 5.3). The ^{1}D term has a considerably longer lifetime than ^{1}S, and therefore will be de-populated at a lower $n_{\rm e}$ by collisional de-excitations, causing a weakening of the λ4959 and λ5007 lines. What starts to play a role

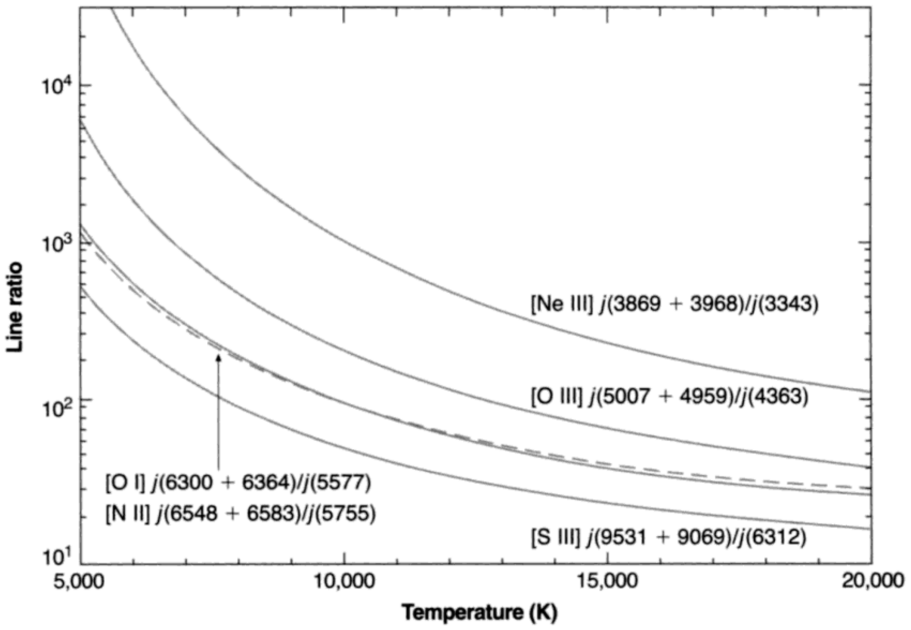


Figure 9.5: *Five temperature sensitive forbidden line ratios as function of the electron temperature. The [O I] (solid line) and [N II] (dashed line) are almost superimposed, partly because they have very similar excitation potentials. All ratios shown are in the limit* $n_{\rm e} \ll n_{\rm e}^{\rm crit}$ *(*$n_{\rm e} = 1\,cm^{-3}$*). From Osterbrock & Ferland,* Astrophysics of Gaseous Nebulae and Active Galactic Nuclei, *2006.*

at electron densities above the value given previously are collisional excitations from ^{1}D to ^{1}S. This strengthens the emission from $\lambda 4363$. A proper description of this problem requires the solution of a more extended set of statistical equilibrium equations, however, an analytical solution that is correct to within first order in $\exp(-\Delta E_{23}/kT)$ is

$$\frac{\eta_{4959} + \eta_{5007}}{\eta_{4363}} = \frac{7.90 \exp(3.29 \times 10^4/T)}{1 + 4.5 \times 10^{-4}\, n_{\rm e}/T^{1/2}} \tag{9.24}$$

In the second edition of Osterbrock & Ferland similar approximations are given for [N III], [Ne III], en [S III]. Note that the correction term for the electron density is very small. Even if only a rough estimate of $n_{\rm e}$ is used, very reasonable estimates for T can be made. Giving Eq. (9.24) some thought, one must admit that it is a fascinating result: the temperature of the nebular gas can be derived without any knowledge of the local radiation field, the distance to the nebula, and (often) the local electron density.

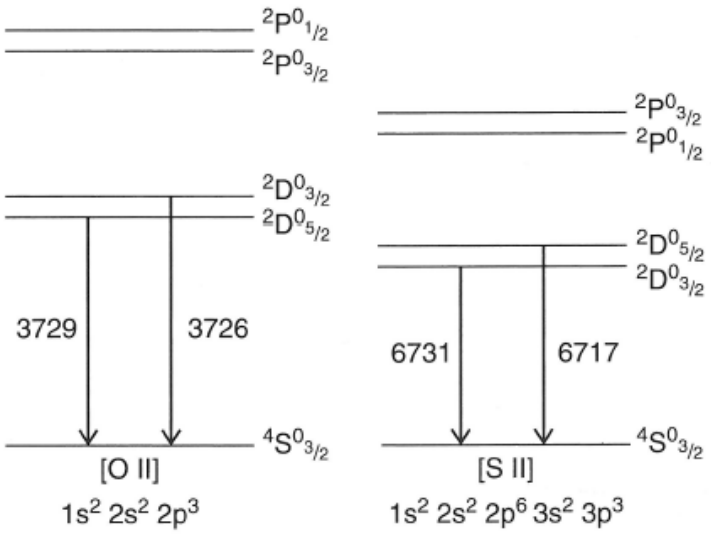


Figure 9.6: *Energy-level diagram of the $2p^3$ ground configuration of O II and $3p^3$ ground configuration of S II, relevant for the formation of forbidden [O II] and [S II] lines. Note that the fine-structure levels $^2D^\circ_{3/2,5/2}$ and $^2P^\circ_{1/2,3/2}$ (the energy separation of the different J-levels is exaggerated for clarity) are switched around in these two ions.*

9.3 n_e diagnostics: collisionally excited forbidden fine-structure lines

An estimate of the electron density n_e in the rarefied nebular gas can be made by determining the ratio of the line strengths of forbidden transitions in the ground configuration of ions with very comparable excitation energy (contrary to what is required for a good T-diagnostic, see Sect. 9.2). The two best examples of such a situation are [O II] $\lambda3729/\lambda3726$ en [S II] $\lambda6716/\lambda6731$[2].

We again take oxygen as an example. To keep a clear view of the situation we refer to the ground level $^4S^\circ_{3/2}$ as level 1, the first excited level $^2D^\circ_{5/2}$ as level 2, and the second excited level $^2D^\circ_{3/2}$ as level 3. The two highest terms in the ground level configuration, $^2P^\circ_{1/2,3/2}$, need not be considered. In formulating the relevant statistical equilibrium equations we consider collisional excitations, collisional de-excitations and spontaneous de-excitations. Note that we ignore the (forbidden) transition $^2D^\circ_{3/2} - {}^2D^\circ_{5/2}$. Though this is a good approximation for the radiative transition between these two fine structure levels, it is not so for the collisional coupling. However, in the two limiting situations $n_e \to 0$ and $n_e \to \infty$ this is not a problem: in the first instance these collisions are indeed negligible; in the second instance the two levels (2 and 3) will be in LTE relative to each other. This is automatically taken care of by considering the collisional transitions between $1-2$ and $1-3$. The situation of arbitrary n_e,

[2]Notice that for the oxygen line the line ratio is longest wavelength / shortest wavelength of the doublet while this is reversed in case of sulpher. This is because in these ions the $^2D^\circ_{3/2,5/2}$ are interchanged, see figure 9.6.

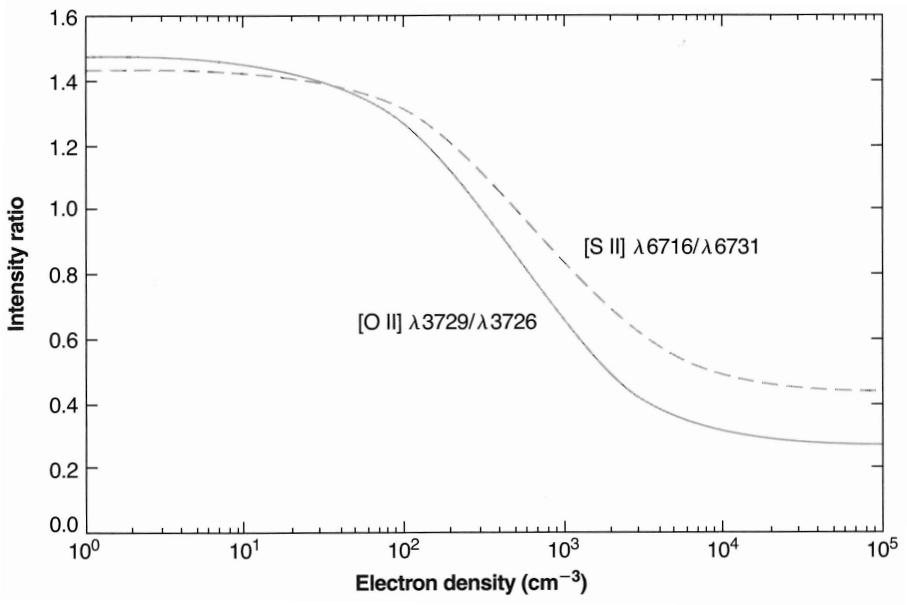


Figure 9.7: *Statistical equilibrium calculation of line ratios for [O II] (solid line) and [S II] (dashed line) as a function of* n_e *for a temperature* $T = 10\,000$ *K. At other temperatures the relations shown are nearly correct if one assumes the horizontal scale represents* $n_\mathrm{e}(10^4/T)^{1/2}$*. From Osterbrock & Ferland,* Astrophysics of Gaseous Nebulae and Active Galactic Nuclei, *2006.*

however, does require a consideration of the $^2\mathrm{D}^\circ_{3/2} - {}^2\mathrm{D}^\circ_{5/2}$ coupling.

Using $C_{ij} = n_\mathrm{e}\, q_{ij}$ for convenience, we get for level 2

$$n_2\,(A_{21} + C_{21}) = n_1\, C_{12}, \tag{9.25}$$

and for level 3

$$n_3\,(A_{31} + C_{31}) = n_1\, C_{13}. \tag{9.26}$$

For the ratio between the line strengths 21 and 31 it follows that

$$\begin{aligned} \frac{\eta_{21}}{\eta_{31}} = \frac{\eta_{3729}}{\eta_{3726}} &= \frac{n_2\, A_{21}\, h\nu_{21}}{n_3\, A_{31}\, h\nu_{31}} \simeq \frac{A_{21}}{A_{31}}\frac{C_{12}}{C_{13}}\frac{(A_{31}+C_{31})}{(A_{21}+C_{21})} = \frac{A_{21}}{A_{31}}\frac{C_{12}}{C_{13}}\frac{C_{31}}{C_{21}}\frac{(A_{31}/C_{31}+1)}{(A_{21}/C_{21}+1)} \\ &\simeq \frac{A_{21}}{A_{31}}\frac{g_2}{g_3}\frac{(A_{31}/C_{31}+1)}{(A_{21}/C_{21}+1)}. \end{aligned} \tag{9.27}$$

The first approximately equal sign denotes that $\nu_{21} \simeq \nu_{31}$. For the last equality we have used that $n_l^\mathrm{LTE}\, C_{lu} = n_u^\mathrm{LTE}\, C_{ul}$, as collisions couple the populations to the local temperature and density, and realized that $E_{12} \simeq E_{13}$.

In the low density limit, each collisional excitation is followed by the emission of a photon. One obtains $\eta_{3729}/\eta_{3726} = C_{12}/C_{13} = \Omega_{12}/\Omega_{13}$. In the high density limit, that for $^2\mathrm{D}^\circ_{5/2}$

is reached at $n_{\rm e}^{\rm crit} \sim 3 \times 10^3\,{\rm cm}^{-3}$ and for $^2{\rm D}_{3/2}$ at $n_{\rm e}^{\rm crit} \sim 1.6 \times 10^4\,{\rm cm}^{-3}$, it follows that $\eta_{3729}/\eta_{3726} = A_{21}\,g_2/A_{31}\,g_3 = 0.34$.

Figure 9.7 shows the behavior of η_{3729}/η_{3726} as function of $n_{\rm e}$ for the exact solution of the statistical equilibrium equations, also accounting for collisional excitations to the $^2{\rm P}^{\circ}_{1/2,3/2}$ levels.

Exercise 9.1

Derive the cooling rate Eq. 9.13 starting from Eq. 6.2.

Exercise 9.2

Observations of an H II region surrounding an O5 Ia supergiant in the Orion Nebula show that the electron temperature of the gas is 9000 K. A similar O5 Ia supergiant in the Small Magellanic Cloud has an H II region with an electron temperature of 11000 K. What could be the cause of the different electron temperatures?

Exercise 9.3

This could be a nice exam question. It is believed that the sun is inside a local bubble of hot gas of about a million degrees and hydrogen density $n_{\rm H} = 10^{-3}\,{\rm cm}^{-3}$. This gas is highly ionized and hence provides no significant opacity for observations within 100 pc or so. However, numerous investigations have shown that there is a minimum column density of neutral hydrogen of about $N_{\rm H} = 10^{18}\,{\rm cm}^{-2}$ toward all stars observed. Therefore it appears that we are embedded in a small, at least partially neutral cloud with approximately this column depth towards its edge. The temperature of the gas in this bubble within a bubble is 8 000 K and the neutral hydrogen density $n_{\rm H}$ = 0.1 cm^{-3}. Assume all H I to be in the ground state. Furthermore, assume that the bubble of neutral gas is spherical and that we are at its center.

a) What is the diameter of this bubble, assuming an atomic hydrogen density of $n_{\rm H} = 0.1\,{\rm cm}^{-3}$.

b) Derive an expression for the optical depth in the Lyα line of hydrogen in terms of the column depth in hydrogen and a constant temperature in the local bubble. Use for the extinction coefficient

$$\chi_\nu = \alpha_{lu}(\nu)\, n_l = \frac{\pi e^2}{m_{\rm e} c}\, f_{lu}\, n_l\, \phi_\nu, \tag{9.28}$$

where $f_{lu} = 0.4162$ and assume a Gaussian line profile function.

c) Is the local bubble optically thin or thick at the center wavelength of Lyα?

d) In spectra of distant quasars we *do* see numerous Lyα lines caused by intervening galactic halos and gas clouds. How can this be?

10

Molecular gas

As we have discussed earlier, molecules are widely used as diagnostic tools in studies of the ISM. Fathoming the chemistry leading to the molecules we observe in space is a fundamental problem in astronomy. It is essential that we grasp the formation, destruction, hence abundance of well-known species such as H_2 and CO. Simple chemical networks may be in operation in e.g. dark clouds, leading to simple organics. Complex organics, either in the ISM, in proto-planetary disks, or in exo-planetary atmospheres might produce building blocks of life as we know them. Observations with ground-based telescopes such as the James Clark Maxwell Telescope (JCMT), the Institut de Radioastronomie Millimétrique (IRAM) interferometer, the Infrared Space Observatory (ISO), the Spitzer Space Telescope (SST) and especially the Herschel Space Observatory (HSO) and Atacama Large Millimeter Array (ALMA) are adding enormously to our knowledge of molecules in space. Today more than 180 molecules are known to exist in the medium in-between the stars, some simple – for instance H_2, some surprisingly complex – for instance *fullerenes*, a family of structures composed entirely of carbon, including spherical, ellipsoidal, or tube like shapes (see Fig. 10.1). Perhaps most intriguing so far is the discovery of the amino acid *glycine* ($C_2H_5NO_2$) – one of the 20 amino acid building blocks of genetic code – in meteorites and possibly in space (but see Snyder et al. 2005, ApJ 619, 914). For an up to date list of molecules in space with literature references, we refer to the websites of the Astrochemistry Laboratory at Goddard Space Flight Center[1] and the University of Köln[2], but one may also check the Wikipedia lemma *List of interstellar and circumstellar molecules*. For an overview of known complex molecules, defined as those that consist of 6 atoms or more, see Fig. 10.2. All contain the element carbon so can be called organic.

[1] `http://science.gsfc.nasa.gov/691/cosmicice/interstellar.html`
[2] `http://www.astro.uni-koeln.de/cdms/molecules`

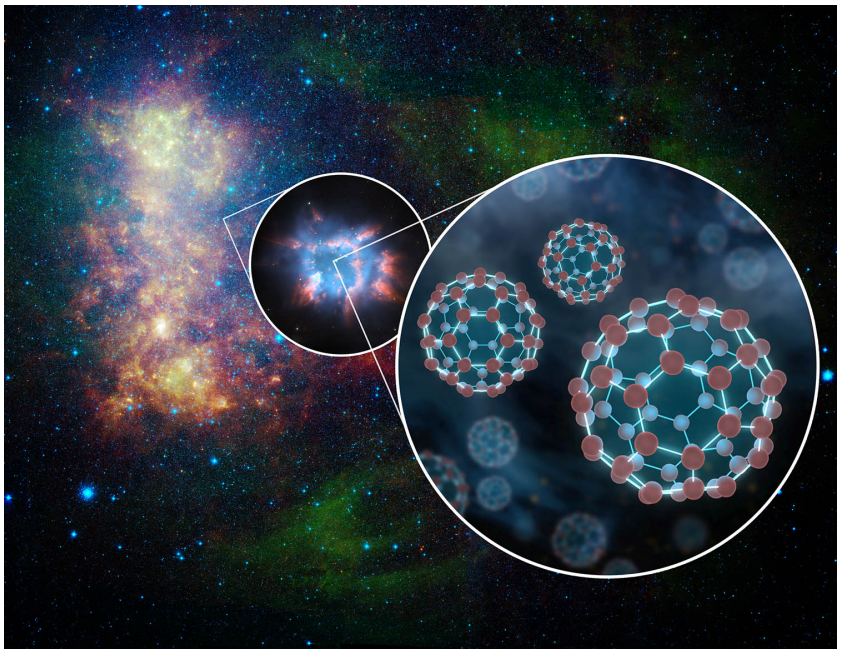


Figure 10.1: *The fullerines C_{60} (buckyballs) and C_{70} have been discovered in a number of Planetary Nebulae, both in the Galaxy and (in once case) in the Small Magellanic Cloud. The figure shows an artist impression of these fullerenes, in the Planetary Nebula SMC 16 in the SMC.*

10.1 Gas phase reactions

Consider the two-body reaction

$$\mathrm{A} + \mathrm{B} \xrightarrow{k} \mathrm{C}, \tag{10.1}$$

where reactants A and B interact and form the product C. The rate at which the abundance of species C increases with time is given by the *rate equation*

$$\frac{\mathrm{d}n(\mathrm{C})}{\mathrm{d}t} = k\, n(\mathrm{A})\, n(\mathrm{B}), \tag{10.2}$$

with $n(\mathrm{X})$ the number density of species X in cm^{-3}. The *rate coefficient* k describes how fast or slow the reaction proceeds, i.e. the probability that the product is formed when the reactants meet. Because the rate of change in the abundance of species C is proportional to the product of the abundances of its two reactants, it is a second order rate equation. In this case, the unit of k is cm^3 s^{-1}.

In case of a one-body reaction, e.g. photo-dissociation,

$$\mathrm{AB} + h\nu \xrightarrow{k} \mathrm{A} + \mathrm{B}, \tag{10.3}$$

Species	Name	Source	Species	Name	Source
Hydrocarbons			**N-Containing**		
C_2H_4	Ethene	circ	CH_3CN	Acetonitrile	cc, hc, of
HC_4H	Butadiyne	circ	CH_3NC	Methylisocyanide	hc
H_2C_4	Butatrienylidene	circ, cc, lc	CH_2CNH	Keteneimine	hc
C_5H	Pentadiynyl	circ, cc	HC_3NH^+	Prot. cyanoacetylene	cc
CH_3C_2H	Propyne	cc, lc	C_5N	Cyanobutadiynyl	circ, cc
C_6H	Hexatriynyl	circ, cc, lc	HC_4N	Cyanopropynylidene	circ
C_6H^-	Hexatriynyl ion	circ, cc, lc	CH_3NH_2	Methylamine	hc, gc
H_2C_6	Hexapentaenylidene	circ, cc, lc	C_2H_3CN	Vinylcyanide	cc, hc
HC_6H	Triacetylene	circ	HC_5N	Cyanodiacetylene	circ, cc
C_7H	Heptatriynyl	circ, cc	CH_3C_3N	Methylcyanoacetylene	cc
CH_3C_4H	Methyldiacetylene	cc	CH_2CCHCN	Cyanoallene	cc
CH_3CHCH_2	Propylene	cc	NH_2CH_2CN	Aminoacetonitrile	hc
C_8H	Octatetraynyl	circ, cc	HC_7N	Cyanotriacetylene	circ, cc
C_8H^-	Octatetraynyl ion	circ, cc	C_2H_5CN	Propionitrile	hc
CH_3C_6H	Methyltriacetylene	cc	CH_3C_5N	Methylcyanodiacetylene	cc
C_6H_6	Benzene	circ	HC_9N	Cyanotetraacetylene	circ, cc
O-Containing			C_3H_7CN	N-propyl cyanide	hc
CH_3OH	Methanol	cc, hc, gc, of	$HC_{11}N$	Cyanopentaacetylene	circ, cc
HC_2CHO	Propynal	hc, gc	**S-Containing**		
c-C_3H_2O	Cyclopropenone	gc	CH_3SH	Methyl mercaptan	hc
CH_3CHO	Acetaldehyde	cc, hc, gc	**N,O-Containing**		
C_2H_3OH	Vinyl alcohol	hc	NH_2CHO	Formamide	hc
c-CH_2OCH_2	Ethylene oxide	hc, gc	CH_3CONH_2	Acetamide	hc, gc
$HCOOCH_3$	Methyl formate	hc, gc, of			
CH_3COOH	Acetic acid	hc, gc			
$HOCH_2CHO$	Glycolaldehyde	hc, gc			
C_2H_3CHO	Propenal	hc, gc			
C_2H_5OH	Ethanol	hc, of			
CH_3OCH_3	Methyl ether	hc, gc			
CH_3COCH_3	Acetone	hc			
$HOCH_2CH_2OH$	Ethylene glycol	hc, gc			
C_2H_5CHO	Propanal	hc, gc			
$HCOOC_2H_5$	Ethyl formate	hc			

Figure 10.2: *Complex organic interstellar molecules. Here complex implies* ≥ 6 *atoms. Abbreviations refer to the location where these species are found: circ = circumstellar envelope around evolved stars or proto-planetary nebulae; cc = cold cloud core; hc = hot cores or hot corinos (a hot corino is the low-mass version of a hot core); lc = lukewarm corino; gc = galactic center cloud; of = outflow. From: Herbst & van Dishoeck 2009, ARAA 47, 427.*

the rate of formation of species A and B are given by a first order rate equation,

$$\frac{\mathrm{d}n(\mathrm{A})}{\mathrm{d}t} = \frac{\mathrm{d}n(\mathrm{B})}{\mathrm{d}t} = k\, n(\mathrm{AB}), \tag{10.4}$$

where the rate of change in the abundance of the product(s) depends only on the abundance of its single reactant. In this case, the unit of k is s^{-1}.

Assuming a species i may form or get destroyed in a set of either one-body or two-body

reactions, it holds that

$$\frac{\mathrm{d}n_i}{\mathrm{d}t} = \sum_{j,k} k_{j,k}\, n_j\, n_k + \sum_l k_l\, n_l - n_i \left(\sum_m k_{i,m}\, n_m + \sum_n k_n \right). \tag{10.5}$$

The first two terms of this ordinary differential equation (ODE) represent all reactions forming species i, either through two-body reactions (first term) or single-body reactions (second term). Equivalently, the final two terms represent all reactions destroying species i, either in two-body reactions (third term) or one-body reactions (fourth term).

The situation in which all species in the network of chemical reactions are present in abundances which have no further tendency to change with time – i.e. the system is in *steady state* – is referred to as *chemical equilibrium*. In case the network consists of a single chemical reaction, this state results when the forward reaction proceeds at the same rate as the reverse reaction. More general, chemical equilibrium for a set of species X_i implies that

$$\frac{\mathrm{d}n(\mathrm{X}_i)}{\mathrm{d}t} = 0 \qquad \forall\, \mathrm{X}_i. \tag{10.6}$$

Photodissociation is only one type of gas-phase reaction. An overview of the different types of gas-phase reactions is given in Table 10.1. They are grouped according to their effect on the species involved. Reactions can lead to bond formation, bond destruction, bond rearrangement, and ionization/neutralization. They can also be divided in two-body reactions, where two species meet for the reaction to occur (Eq. 10.1), and one-body reactions, where a species is photo-ionized or photo-dissociated by either a UV photon or a cosmic-ray (e.g. Eq. 10.3).

Two-body reactions

The temperature dependence of the reaction rate coefficient is described by the Arrhenius equation, proposed by Svante Arrhenius (1859 - 1927) based on the work of the Dutch chemist Jacobus van 't Hoff (1852 - 1911),

$$k = \alpha \left(\frac{T}{300} \right)^{\beta} \exp\left(-\frac{\gamma}{T} \right). \tag{10.7}$$

The coefficient α is the pre-exponential factor; β specifies the temperature dependence. The energy barrier of the reaction, i.e. the energy that needs to be available to the system of reactants for the reaction to occur, is captured with the parameter $\gamma = E_a/k$, where E_a is the activation energy in K.

The value of the rate coefficient of a two-body reaction and its temperature dependence depend mainly on the charged state of the reactants. Reactions between neutral molecules are generally slow, since these reactions have energy barriers due to the electronic rearrangement required to form the products. Ion-neutral reactions are faster than neutral-neutral reactions, thanks to the long-range interaction between the ion and the electric dipole of the neutral

Bond formation reactions		Typical k	Unit
Radiative association	$X^{(+)} + Y \rightarrow XY^{(+)} + h\nu$	$10^{-17} - 10^{-14}$	$cm^3\,s^{-1}$
Associative detachment	$X^- + Y \rightarrow XY + e^-$	$\sim 10^{-9}$	$cm^3\,s^{-1}$
Bond destruction reactions			
Collisional dissociation*	$XY + Z \rightarrow X + Y + Z$	$10^{-9} - 10^{-7}$	$cm^3\,s^{-1}$
Dissociative recombination	$XY^+ + e^- \rightarrow X + Y$	$10^{-7} - 10^{-6}$	$cm^3\,s^{-1}$
Photo-dissociation	$XY + h\nu \rightarrow X + Y$	$10^{-11} - 10^{-8}$	s^{-1}
Cosmic-ray induced dissociation	$XY + cr \rightarrow X + Y$	1.3×10^{-17}	s^{-1}
Bond rearrangement reactions			
Ion-neutral	$X^{+,-} + YZ \rightarrow XY^{+,-} + Z$	$10^{10} - 10^{-8}$	$cm^3\,s^{-1}$
Neutral-neutral	$X + YZ \rightarrow XY + Z$	$10^{-13} - 10^{-10}$	$cm^3\,s^{-1}$
Charge exchange	$X^+ + YZ \rightarrow X + YZ^+$	$\sim 10^{-9}$	$cm^3\,s^{-1}$
Mutual neutralization	$X^- + Y^+ \rightarrow X + Y$	$\sim 10^{-8}$	$cm^3\,s^{-1}$
Ionization/neutralization reactions			
Radiative recombination	$X^+ + e^- \rightarrow X + h\nu$	$\sim 10^{-12}$	$cm^3\,s^{-1}$
Radiative electron attachement	$X + e^- \rightarrow X^- + h\nu$	$10^{-9} - 10^{-7}$	$cm^3\,s^{-1}$
Photoionization	$X + h\nu \rightarrow X^+ + e^-$	$10^{-11} - 10^{-8}$	s^{-1}
Cosmic-ray induced ionization	$X + cr \rightarrow X^+ + e^-$	$\sim 1.3 \times 10^{-17}$	s^{-1}

Table 10.1: *Types of chemical reactions and their typical rate coefficients k. From McElroy et al. (2013) and van Dishoeck (2014). cr stands for cosimic ray. *Also known as thermal dissociation.*

molecules, which is either a permanent dipole in the case of polar molecules, or an induced dipole moment in case of non-polar molecules. Mutual neutralization and dissociative recombination reactions, involving a positively charged and a negatively charged reactant, are even faster than ion-neutral reactions, as the long-range attraction between the reactants is stronger.

One-body reactions

The rate coefficient of a one-body reaction depends on the specific process. Cosmic rays with energies between 10 and 100 MeV may cause direct ionization at a rate

$$k = \zeta. \tag{10.8}$$

For instance, for direct cosmic-ray ionization of H_2, $\zeta = 1.36 \times 10^{-17}\,s^{-1}$. In this process secondary electrons with a mean energy around 30 eV are produced that in turn may excite, dissociate, and ionize other molecules (mainly H_2). Excitations of H_2 to the B $^1\Sigma_u^+$, C $^1\Pi_u$ (see Sect. 4.2 for the designation of molecular energy levels) and higher electronic states of H_2 are followed by spontaneous emission of photons which have sufficient energy to dissociate many of the interstellar molecules. The rate of cosmic-ray-induced photo-reactions these (UV) photons cause is

$$k = \zeta \left(\frac{T}{300}\right)^{\beta} \frac{\gamma}{1-\omega}, \tag{10.9}$$

where ζ is the cosmic-ray ionization rate, γ is the efficiency of the cosmic-ray ionization event (as defined in Eq. 8 of Gredel et al. 1989), and ω is the dust-grain albedo in the far ultraviolet (typically 0.4−0.6 at 150 nm).

The photo-dissociation or photo-ionization rate of a species upon interaction with a UV photon is calculated as (see Sect. 8.2)

$$k = 4\pi \int_{\nu_\circ}^{\infty} \alpha_\nu^{\rm bf} \frac{J_\nu}{h\nu}\, d\nu = c \int_{\nu_\circ}^{\infty} \alpha_\nu^{\rm bf} \frac{u_\nu}{h\nu}\, d\nu, \tag{10.10}$$

where $\alpha_\nu^{\rm bf}$ is the photo-dissociation or photo-ionization cross section and u_ν the energy density of the interstellar radiation field (see Eq. 6.8).

10.2 Formation of H_2

For many years it was a mystery as to how molecules could form in the low-density environments of the interstellar medium in molecular clouds. The basic problem is that when two atoms collide in the gas phase the most likely outcome is that they will simply bounce off. A third body is normally needed to carry away the excess (binding) energy that is liberated when two atoms bind (see Sect. 4.1).

The H_2 molecule exemplifies this problem. When two free H atoms, both in the ground electronic state, approach one another, by symmetry there is no electric dipole moment. Consequently, there is no electric dipole radiation that could carry away energy from the system and leave the two H atoms in a bound state. Electric quadrupole transitions are possible, and in principle allow for the radiative association reaction

$$\mathrm{H} + \mathrm{H} \rightarrow \mathrm{H_2} + h\nu \tag{10.11}$$

to occur, but the rate at which this happens is so low that it can be ignored in astrochemistry. The three-body reaction

$$\mathrm{H} + \mathrm{H} + \mathrm{H} \rightarrow \mathrm{H_2} + \mathrm{H} + E_{\rm kin}, \tag{10.12}$$

in which the H atom is carrying off the 4.48 eV of energy released when H_2 is formed, too has a reaction rate that is so low that it is negligible at interstellar or intergalactic densities.

So, how does H_2 form?

Gas-phase formation of H_2

There is actually no good way to produce molecular hydrogen through gas phase reactions, and in the present-day universe H_2 is formed on interstellar grain surfaces (see below). However, in the early universe, when dust was absent, gas-phase formation may have played an important role prior to heavy element synthesis in the first generation of stars.

The dominant channel for H_2 formation in the gas phase starts by the formation of the hydrogen anion H^- by radiative association

$$H + e^- \rightarrow H^- + h\nu, \tag{10.13}$$

of which the *rate coefficient* or *rate constant* $k \sim 4.0 \times 10^{-16}\,(T/300\,\mathrm{K})^{2/3}\,\mathrm{cm^3\,s^{-1}}$. Hence, the number of H^- that is produced per $\mathrm{cm^3}$ per second is $\mathrm{d}n(H^-)/\mathrm{d}t = k\,n(H)\,n_e$. Subsequently, H_2 is formed by associative detachment

$$H^- + H \rightarrow H_2(v, J) + e^- + E_{\mathrm{kin}}, \tag{10.14}$$

of which the rate coefficient $k \sim 1.3 \times 10^{-9}\,\mathrm{cm^3\,s^{-1}}$. In the diffuse ISM, most of the H^- that is formed by reaction (10.13) is actually not available for the follow-up reaction leading to molecular hydrogen, but is destroyed by the photo-dissociation reaction

$$H^- + h\nu \rightarrow H + e^-, \tag{10.15}$$

of which the rate coefficient $k \sim 2.4 \times 10^{-7}\,\mathrm{s^{-1}}$ assuming an average ambient interstellar radiation field, resulting in a very low formation rate of H_2. In addition, H^- can also be destroyed by mutual neutralization reactions with protons ($H^- + H^+ \rightarrow H + H$) or other positive ions.

Grain surface catalysis of H_2

Because of the difficulty of forming H_2 in the gas phase, it is now generally accepted that the formation of H_2 in the ISM proceeds via *grain catalysis*. The idea is that H atoms adsorb to the surface of a dust grain, i.e. H atoms arrive and become bound to the grain surface, after which they migrate over that surface. During the migration process, part of the H atom population may arrive at a site where it is bound strongly enough to (at least temporarily) prevent it from being evaporated by thermal fluctuations on the grain surface. It then has to wait until it is encountered by a newly arrived H atom, wandering over the grain surface, and form H_2.

The formation rate of H_2 in $\mathrm{cm^{-3}\,s^{-1}}$ can be expressed as

$$\frac{\mathrm{d}n(H_2)}{\mathrm{d}t} = \frac{1}{2}\,S(T, T_\mathrm{d})\,\eta(T_\mathrm{d})\,\sigma_\mathrm{d}\,n_\mathrm{d}\,n_\mathrm{H}\,v_\mathrm{H}, \tag{10.16}$$

where $S(T, T_\mathrm{d})$ is the sticking probability of an H atom with temperature T colliding with a grain of temperature T_d, η is the probability that an adsorbed H atom will migrate over the grain surface, find another H atom, and form H_2 before evaporating from the grain surface. The dust grain density is n_d and $\sigma_\mathrm{d} = \pi\,a_\mathrm{d}^2$ the grain geometric cross section, where a_d is the grain size. n_H is the H atom density, and v_H is the thermal speed of the H atoms, given by Eq. (2.62). The factor $\frac{1}{2}$ enters because it takes two H atoms to form an H_2 molecule. The energy released when two H atoms react to form H_2 in the ground state is $\Delta E = 4.5\,\mathrm{eV}$. This energy is large enough to overcome the forces that were binding the two free H atoms to the grain, and the H_2 molecule is ejected from the grain surface.

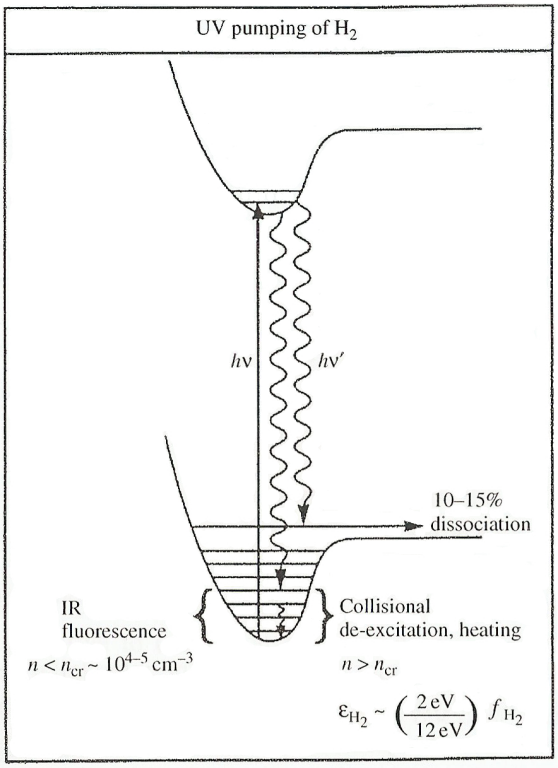


Figure 10.3: *A schematic diagram illustrating the H_2 pumping, fluorescence, dissociation, and heating through the absorption of far-UV photons. Absorption of a far-UV photon followed by FUV radiative decay can leave H_2 vibrationally excited in the ground vibrational state. The excess vibrational energy can be emitted as a near-IR photon or the molecule can be de-excited through collisions, thereby heating the gas. In about 10-15% of the far-UV pumps, H_2 decays to the vibrational continuum of the ground electronic state and the molecule dissociates. From: Tielens (2005).*

S and η tend to decrease with increasing temperature as atoms bounce from the grain surfaces or evaporate before reacting more easily. For sufficiently low temperatures, $S \sim 1$ and $\eta \sim 1$. The larger the grains in the population, the slower the formation rate, as the total geometric grain cross section $n_{\rm d}\,\sigma_{\rm d}$ becomes less. For a typical grain radius of $0.1\,\mu$m (or $10^{-5}\,$cm) we find $\sigma_{\rm d} = 3 \times 10^{-10}$. Using Eq.5.3 we obtain for $n_{\rm d}/n_{\rm H} \sim 4 \times 10^{-12}$, such that $\sigma_{\rm d} \sim 10^{-21}$ cm^2 per H atom, but it could conceivably be much larger if a large population of very small $a_{\rm d} \lesssim 50$ Å grains is present. Models that, in addition to silicates and graphite particles also include PAHs typically have a $\gtrsim 5$ times higher total grain cross section. However, the application of these models to observations seems to indicate that for PAHs the value of η is very low.

Though many of the details need sorting out, it appears that with the canonical numbers given here, molecular hydrogen formation on interstellar grain surfaces can explain quantitatively the observed H_2 abundances.

10.3 Destruction of H_2

The destruction of molecular hydrogen is dominated by photo-dissociation processes.

10.3.1 Photo-dissociation of H_2

Photo-dissociation is the principal process destroying interstellar H_2. It is quite an efficient process and in the absence of self-shielding, diffuse H I clouds will contain only trace amounts of H_2. The outcome of photo-dissociation is that

$$H_2 + h\nu \rightarrow H + H + E_{\rm kin}. \tag{10.17}$$

The first step in the H_2 photo-dissociation is the absorption of a resonance line photon at far-UV wavelengths with photon energy in the range of 11.2 to 13.6 eV, raising the H_2 from an initial level $X(v, J)$ of the ground electronic state X $^1\Sigma_g^+$ to a level $B(v', J')$ or $C(v', J')$ of the first or second electronic excited states, B $^1\Sigma_u^+$ and C $^1\Pi_u$ (see Sect. 4.2 for the designation of molecular energy levels). Photons leading up to excitation of the B $^1\Sigma_u^+$ electronic state are referred to as Lyman(-band) photons; those leading up to excitation of the C $^1\Pi_u$ state as Werner(-band) photons. The original photo-excitation is via a permitted absorption line, and therefore the newly excited level $B(v', J')$ or $C(v', J')$ will have electric dipole-allowed decay channels. In about $\sim$ 85-90% of cases, the exited level decays to vibrationally excited bound levels $X(v'', J'')$ of the ground electronic state. Sometimes, however, spontaneous decay of the excited level $B(v', J')$ or $C(v', J')$ will be to the *vibrational continuum* of the ground electronic state: the molecule will 'vibrate' apart, separating into two free H atoms (see Figs. 10.3 and 4.10).

10.4 Structure of a photo-dissociation region or PDR

In high density regions the gas is able to shield itself against radiation that is capable of destroying molecules by photo-dissociation. This *self-shielding* refers to the phenomenon where the photo-excitation transitions become optically thick, so that the molecule in question is "shielded" from starlight by other identical molecules.

Because a high density is required, molecular gas is often found in star-forming regions – indeed, stars form out of molecular gas. When a massive star forms, it may strongly irradiate the remaining molecular cloud material with ultraviolet radiation, resulting in photo-dissociation and photo-ionization. The photo-ionized gas, heated to $\sim 10^4$ K, will be over-pressured and drive a compressive wave (possibly a shock wave) in the molecular cloud, i.e. the ionized gas will try to flow toward lower-pressure regions nearby (see Sect. 8.6).

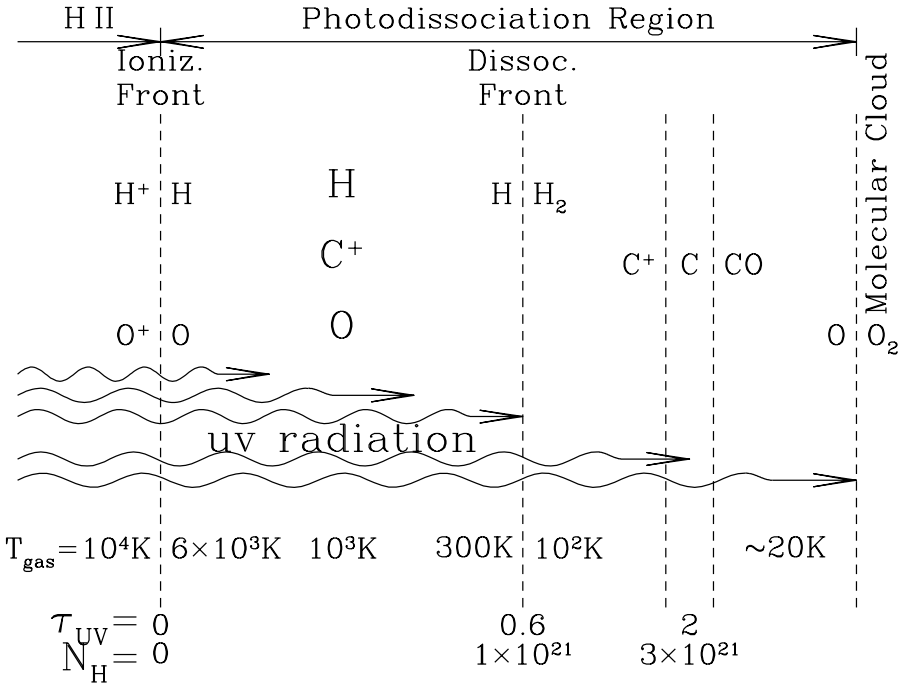


Figure 10.4: *Structure of a PDR at the interface between an H II region and a dense molecular cloud. The front is assumed to be illuminated from the left by unidirectional radiation. Measured from the ionization front, the hydrogen column density and optical depth increase to the right. Dissociation fronts of molecules other than H_2 will also be present. From: Drain (2011).*

The interface between the H II region and the dense molecular cloud is called a *photo-dissociation region* or *photon-dominated region* or PDR, and its structure is illustrated in Fig. 10.4. The PDR is bounded by an *ionization front* – the surface where the hydrogen is half ionized and half neutral – and will contain a *photo-dissociation front* – the surface where the hydrogen is half atomic and half molecular (by mass). If we adopt a frame of reference in which the photo-dissociation front is at rest, then the molecular gas will flow toward the photo-dissociation front where it is dissociated, after which the atomic gas flows away from the photo-dissociation front toward the ionization front.

Diffuse molecular clouds have a qualitatively similar structure, although they may lack both the hottest and the coolest regions shown in Fig. 10.4, depending on whether they are bounded by photo-ionized gas, on the one hand, and how thick they are, on the other hand. Figure 10.5 shows the profile of a model plane-parallel cloud with an H I / H_2 transition. The cloud is illuminated from one side by unidirectional radiation with the energy density and spectrum of the interstellar radiation field (ISRF). The gas is assumed to be at uniform pressure $p/k = 3000\,\mathrm{cm}^{-3}\,\mathrm{K}$, i.e. the entire medium is in pressure equilibrium. The cosmic ray ionization

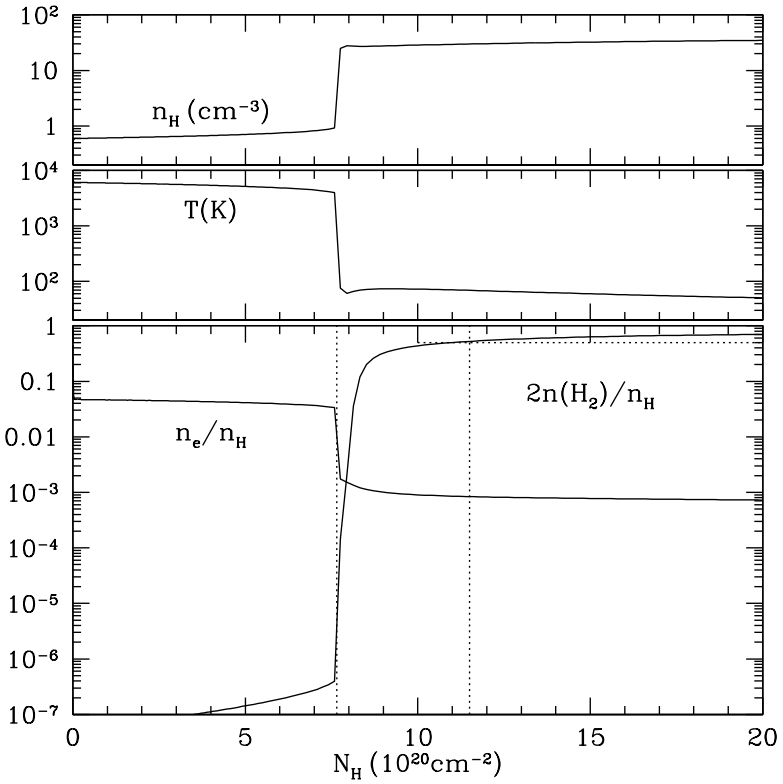


Figure 10.5: *Profile of the H I−H_2 transition in a diffuse molecular cloud, for an assumed pressure $p/k = 3000$ cm^{-3} K. The dotted lines delimit the surface layer of cool gas where more than half of the hydrogen is H I. From: Drain (2011).*

rate $\xi_{\rm CR} = 2 \times 10^{-16}\,{\rm s}^{-1}$ and the properties of the dust that is present and that is attenuating starlight, causing photo-electric heating and formation of H_2, is standard. The gas is further assumed to be in thermal and chemical equilibrium at each point, with heating = cooling, ionization = recombination, and H_2 destruction = H_2 formation. This may be expected to be valid if the flow velocities referred to above are sufficiently small. As the radiation field entering from the left is attenuated by dust, the gas makes a transition from the warm (WNM) phase to the cool (CNM) phase.

In the WNM (on the left of the figure), the H_2 abundance is very low, $\lesssim 10^{-6}$. The steady state H_2 abundance rises as one enters the CNM, as a result of both

- the increased gas density – promoting H_2 formation, and
- growing self-shielding – lowering the photo-dissociation rate.

The onset of the zone where the gas forms molecules has $N_{\rm H} = 3.9 \times 10^{20}$ cm^{-2}, and a dust column with $E(B-V) = 0.066$ mag (see Eq. 12.19), $A_V = 3.1E(B-V) = 0.2$ mag. The optical depth τ_V in dust is here about 0.2 (see Eq. 12.13). The end of this zone is at $N_{\rm H} = 11.5 \times 10^{20}$, i.e. an $\tau_V = 0.6$. The H_2 in the cloud is undergoing far-UV pumping which results in destruction of the H_2 in $\sim$ 10-15 % of the time; the remaining $\sim$ 85-90 % of the UV excitations create a population of rotationally excited H_2 in the cloud.

Orion's Bar

The Orion Bar (see Fig. 10.6) is the most famous example of a PDR. Moving outward from θ^1 Ori C, it includes a high pressure layer of photo-ionized gas, an ionization front, a photo-dissociation zone where the hydrogen is neutral but primarily atomic, and a photo-dissociation front. Tielens et al. (1993) provide a nice overview of the structure of the Orion Bar, in which many different molecules have been detected.

Star-forming galaxies

In star-forming galaxies, an appreciable fraction ($\sim$10 %) of the total luminosity of the galaxy is reprocessed through dense PDRs at the interface between molecular clouds and H II regions. Energy, originally radiated by hot stars is absorbed by molecules and dust grains in the PDR, and re-radiated at longer wavelengths as IR emission from dust and PAHs, and line emission from atoms and molecules in the gas. Part of the starlight energy goes into changing the physical state of the gas – from cold and molecular to hot, photo-dissociated, and possible photo-ionized if an ionization front is present.

10.5 Chemistry in molecular clouds

Once H_2 has formed, other chemistry can follow (see Draine 2011). For other molecules than H_2 the general ultraviolet background provided by starlight too is lethal, with either photo-dissociation or photo-ionization occurring rapidly. In the diffuse ISM, small molecules have photo-dissociation rates that range from $\sim 4\times10^{-11}$ s^{-1} (e.g. H_2) to $\sim 10^{-9}$ s^{-1} (e.g. CN and H_2CO). In clouds, the ultraviolet radiation is attenuated by dust, and the photo-dissociation rates are reduced. The amount by which it is reduced depends on the overall column density (or A_V) through the cloud, and the location within the cloud. For a Giant Molecular Cloud (GMC) the total thickness is typically $A_V \sim 10$. Already at $A_V \sim 3$ the photo-dissociation rate can be reduced by factors of $\sim 10^3 - 10^5$, rendering unimportant photo-dissociation by photons originating *outside* the cloud. However, cosmic rays penetrating the cloud not only ionize H_2 and He, they also cause electronic excitation of H_2 by two processes. First, the electric field of passing charged cosmic rays can excite electrons to bound states (e.g. the B $^1\Sigma_u^+$ and C $^1\Pi_u$ states of H_2) followed by spontaneous emission of an UV photon. Second,

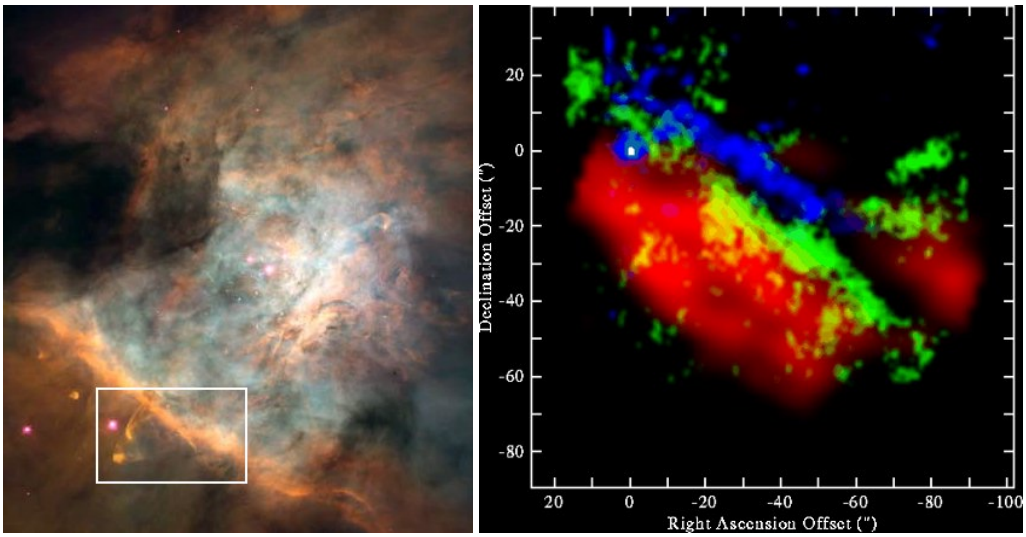


Figure 10.6: *Left panel: In the white box is part of Orion's Bar, the PDR in the Orion nebula. It is here that the ionization of the Trapezium stars is eating its way into the parent molecular cloud. We observe the Bar edge-on, which is a favorable angle to study the individual layers that make up the structure of this PDR. Moving to the south-east (down and left), we pas from the ionized medium into a (swept-up) dusty interface, into the molecular region. Right panel: the PAH emission layer is code in blue, the H_2 emission in green, and the CO emission in red. The yellow color results form an overlap of green and red. From: Tielens et al. 1993.*

the secondary electrons produced by cosmic ray ionization events can themselves excite electronic states of H_2. Together, these two processes result in the generation of UV photons at a rate proportional to the cosmic ray ionization rate.

With H_2 present and UV radiation present, the chemical network in Molecular Clouds (with a typical temperature $T < 100\,\mathrm{K}$) is dominated by *ion-molecule reactions*. Things start with the cosmic-ray induced ionization reaction (see e.g. Ward-Thompson & Whitworth 2011)

$$\mathrm{H_2} + \mathrm{cr} \rightarrow \mathrm{H_2^+} + e^- + \mathrm{cr}, \tag{10.18}$$

where we have denoted the cosmic ray as cr. The ionization helps to drive the subsequent chemistry. This leads to

$$\mathrm{H_2^+} + \mathrm{H_2} \rightarrow \mathrm{H_3^+} + \mathrm{H}, \tag{10.19}$$

producing the highly reactive molecular ion H_3^+. Protonated molecular hydrogen, H_3^+, is the simplest polyatomic molecule. It is the most abundantly produced interstellar molecule, next only to H_2, although its steady state abundance is low because of its extremely high chemical reactivity. H_3^+ is a strong acid (proton donor) and initiates chains of ion-molecule reactions leading to formation of complex molecules (Oka 2006, PNAS 103, 33).

Formation and destruction of OH

H_3^+ can drive the oxygen chemistry, leading to the formation of the water molecule, H_2O, and the OH radical by the following direct route

$$O + H_3^+ \rightarrow OH^+ + H_2. \tag{10.20}$$

An alternative formation mechanism for OH^+ could be by means of ionized oxygen, using

$$H_2 + O^+ \rightarrow OH^+ + H. \tag{10.21}$$

Then

$$OH^+ + H_2 \rightarrow H_2O^+ + H \tag{10.22}$$

followed by

$$H_2O^+ + e^- \rightarrow \begin{cases} OH + H & (20\%) \\ O + H_2 & (9\%) \\ O + H + H & (71\%), \end{cases} \tag{10.23}$$

and

$$H_2O^+ + H_2 \rightarrow H_3O^+ + H, \tag{10.24}$$

followed by

$$H_3O^+ + e^- \rightarrow \begin{cases} O + H_2 + H & (1\%) \\ OH + H_2 & (14\%) \\ OH + H + H & (60\%) \\ H_2O + H & (25\%), \end{cases} \tag{10.25}$$

The OH that is formed in this way is destroyed primarily by photo-dissociation, i.e.

$$OH + h\nu \rightarrow O + H. \tag{10.26}$$

Formation and destruction of CO

In diffuse molecular clouds, most of the gas-phase carbon is in the form of C^+. The formation of CO starts with the (slow) reaction

$$C^+ + H_2 \rightarrow CH_2^+ + h\nu, \tag{10.27}$$

where we have denoted the emitted photon by $h\nu$. The radical CH_2^+ that is produced reacts rapidly with electrons

$$CH_2^+ + e^- \rightarrow \begin{cases} CH + H & (25\%) \\ C + H_2 & (12\%) \\ C + H + H & (63\%), \end{cases} \tag{10.28}$$

producing CH about 25% of the time. The CH produced by dissociative recombination of CH_2^+ can then react with O to produce CO via reaction

$$CH + O \rightarrow CO + H. \tag{10.29}$$

Equation (10.29) is an example of an exchange reaction. Other pathways to forming CO also appear to involve such exchange reactions. They are

$$C + OH \rightarrow CO + H, \tag{10.30}$$

or

$$C^+ + OH \rightarrow CO^+ + H. \tag{10.31}$$

The latter process proceeds to CO via

$$CO^+ + H_2 \rightarrow HCO^+ + H, \tag{10.32}$$

and

$$HCO^+ + e^- \rightarrow CO + H. \tag{10.33}$$

Aside from the initial production of H_2 via grain catalysis, it is assumed that all other reactions resulting in formation of CO in diffuse molecular clouds take place in the gas phase. Whether this is actually the case is uncertain. For example, one could imagine that C and O atoms might stick to silicate grains and react to form CO, with the CO molecule returned to the gas phase either by the energy released in formation of CO, or by photo-desorption. Like OH, CO is destroyed primarily by photo-dissociation.

(1)	$\Rrightarrow$	Ar + cr	$\xrightarrow{k_1}$	$Ar^+ + e^-$	$k_1 = \zeta(\mathrm{Ar}) = 11.4\,\zeta(\mathrm{H})$
(2)		$Ar + H_2^+$	$\xrightarrow{k_2}$	$Ar^+ + H_2$	
(3)		$Ar + H_3^+$	$\xrightarrow{k_3}$	$ArH^+ + H_2$	
(4)		$Ar^+ + e^-$	$\xrightarrow{k_4}$	$Ar + h\nu$	
(5)	$\Rrightarrow$	$Ar^+ + H_2$	$\xrightarrow{k_5}$	$ArH^+ + H$	
(6)	$\Rrightarrow$	$ArH^+ + H_2$	$\xrightarrow{k_6}$	$Ar + H_3^+$	$k_6 = 8.0 \times 10^{-10}\,\mathrm{cm^3\ s^{-1}}$
(7)		$ArH^+ + CO$	$\xrightarrow{k_7}$	$Ar + HCO^+$	
(8)	$\Rrightarrow$	$ArH^+ + O$	$\xrightarrow{k_8}$	$Ar + OH^+$	$k_8 = 8.0 \times 10^{-10}\,\mathrm{cm^3\ s^{-1}}$
(9)		$ArH^+ + C$	$\xrightarrow{k_9}$	$Ar + CH^+$	
(10)		$ArH^+ + e^-$	$\xrightarrow{k_{10}}$	Ar + H	
(11)	$\Rrightarrow$	$ArH^+ + h\nu$	$\xrightarrow{k_{11}}$	$Ar^+ + H$	

Table 10.2: *Simple chemical network describing the formation and destruction pathways of* ArH^+. *The cosmic-ray ionization rate of hydrogen* $\zeta(\mathrm{H}) = 2.0 \times 10^{-16}\,s^{-1}$.

Exercise 10.1

Argonium (ArH^+), the first nobel gas molecule to be detected in space (Barlow et al. 2013, Science 342, 1343), is considered a very good tracer of the atomic gas in the ISM (Schilke et al. 2014, A&A 566, A29). Table 10.2 provides a simple chemical network describing the formation and destruction pathways of ArH^+.

a) Classify each type of reaction in the chemical network using Table 10.1.

b) Write the ordinary differential equation which describes the rate of change of the number density $n(ArH^+)$ of argonium as a function of time.

c) Assuming an equilibrium situation, i.e. $dn/dt = 0$ for all species, derive the following expression for the abundance ratio of ArH^+ and Ar in the diffuse ISM

$$\frac{n(\mathrm{ArH^+})}{n(\mathrm{Ar})} = \frac{\zeta(\mathrm{Ar})}{k_6\, n(\mathrm{H_2}) + k_8\, n(\mathrm{O})}, \tag{10.34}$$

where $\zeta(\mathrm{Ar}) = k_1$ is the cosmic-ray ionization rate of Ar.

According to Schilke et al. 2014, the only reactions contributing significantly to the formation and destruction of ArH^+ are the ones indicated with a $\Rrightarrow$ in Table 10.2. Use for the abundance of oxygen in a diffuse interstellar cloud $n(\mathrm{O}) = 2.9 \times 10^{-4}\, n_\mathrm{H}$ and for argon $n(\mathrm{Ar}) = 3.2 \times 10^{-6}\, n_\mathrm{H}$, where n_H is the total hydrogen particle density (so, all hydrogen in atomic and molecular form).

d) Show that

$$n(\mathrm{ArH^+}) = \frac{4.56 \times 10^{-10}}{25 f(\mathrm{H_2}) + 1.45 \times 10^{-2}} \quad \mathrm{cm^{-3}}, \tag{10.35}$$

where $f(\mathrm{H_2}) = 2n(\mathrm{H_2})/n_\mathrm{H}$ is the molecular mass fraction of hydrogen.

e) Make a plot (e.g. using Excel or python) to show the relation between $n(\mathrm{ArH^+})$ and $f(\mathrm{H_2})$. Is argonium a good tracer of molecular gas or of atomic gas?

Exercise 10.2

In a molecular cloud, molecular hydrogen forms on the surface of grains and is destroyed by the interaction with sufficiently energetic photons.

a) Give formulae for $dn(\mathrm{H_2})/dt$ (formation) and $dn(\mathrm{H_2})/dt$ (destruction). Given is that the photo-destruction cross-section of molecular hydrogen due to photons with frequency ν is $\alpha_{H_2}(\nu)$ (see Section 8.2) and that the minimum frequency where a photon can destroy molecular hydrogen is $\nu_\circ$.

b) Compute the equilibrium molecular hydrogen fraction $x = n(\mathrm{H_2})/N$, where $N = n(\mathrm{H_2}) + n_\mathrm{H}$.

Exercise 10.3

This could be a nice exam question. The molecule AB is formed in the following ways:

$$\begin{aligned} \mathrm{A + B} &\rightarrow \mathrm{AB} + h\nu & \mathrm{k_1} \\ \mathrm{A + BC} &\rightarrow \mathrm{AB + C} & \mathrm{k_2} \end{aligned}$$

The rate coefficients for these reactions are k_1 and k_2 cm^3 sec^{-1}, respectively. AB is lost in the reaction

$$\begin{aligned} \mathrm{AB + D} &\rightarrow \mathrm{A + BD} & \mathrm{k_3} \\ \mathrm{AB} + h\nu &\rightarrow \mathrm{A + B} & \mathrm{k_4} \end{aligned}$$

where the latter is a photo-dissociation reaction. The rate coefficients are k_3 cm^3 sec^{-1} and k_4 sec^{-1}.

a) Write down the expression for the equilibrium number density of AB, n_AB, in terms of other number densities. Why do the number densities of species C and BD *not* feature in this expression?

b) We assume that these are the only four chemical reactions that occur in the medium. To what value will n_AB eventually evolve, again expressed in terms of other number densities?

c) We take away the molecules A, B, and D at time $t_\circ$, when the number density of AB is $n_\mathrm{AB}(t_\circ)$. We further set $k_1 = 0$. How does the number density of AB evolve in time from $t_\circ$ on?

11

Interstellar dust

The existence of solid state material between the stars was first proposed by the Russian astronomer Otto Wilhelm von Struve (1819-1905) in 1847 based on the analysis of star counts which suggested that the number of stars per unit volume decreased with increasing distance from the sun. Struve proposed that the starlight was experiencing absorption proportional to distance. It was not until 1909 that Jacobus Cornelius Kapteyn (1851-1922) realized the full significance of this interstellar extinction. Shortly thereafter Edward Barnard documented the irregular variations in the distribution of the absorbing matter. The identification of small solid state particles as the source of this extinction was finally accepted in the 1930s through the work of Trumpler and Stebbins, Huffer, and Whitford. Over the succeeding decades, we have built on these pioneering studies, but many aspects of interstellar dust – including its chemical composition – remain uncertain.

Interstellar dust plays an important role in many processes in galaxies and in the life cycle of stars. Only roughly one percent of the interstellar matter is in the form of dust (see Sect. 5.1). Because of its efficient absorption of stellar light at short wavelengths, and the re-emission of the absorbed energy at long wavelengths, dust plays an important role in the energy budget of the ISM. For instance, it shields molecular clouds from the UV radiation of stars, so that molecules can exist in the ISM. Dust also provides a surface on which various chemical and physical processes can occur. Dust also plays a crucial role in the process of star formation in contracting molecular clouds, and is the building block from which planets form.

Asymptotic Giant Stars and Red Supergiants produce copious amounts of dust in their stellar winds, so-called *circumstellar dust*. This dust is blown into the ISM and mixes with the material already there. Supernova explosions can also produce dust; in fact they are potentially a very important source of interstellar dust, but there is still considerable uncertainty about the amount of dust these explosions produce. Finally, we mention so-called *ultraluminous infrared galaxies* at redshifts mostly between $z = 1$ to 3. These are distant galaxies in which more than 90 percent of the light is emitted at far-IR wavelengths (see Fig. 11.1). They obviously have large amounts of dust, probably due to a recent starburst. The dust in these galaxies is heated by the many young hot stars that formed during the starburst. Clearly these systems can best be studied at long, infrared and millimeter wavelengths.

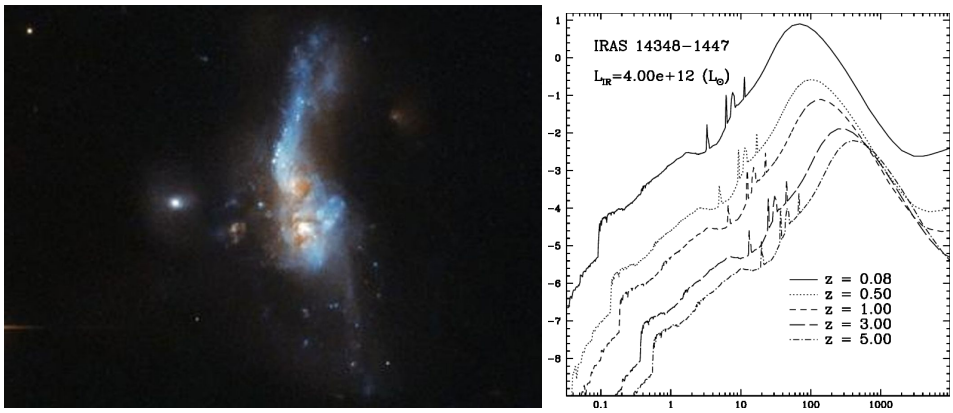


Figure 11.1: Left*: HST/ADS image IRAS 14348-1447, a galaxy merger event of two gas-rich spiral galaxies. Almost 95% of the energy emitted by the system is in the far-infrared. At a redshift of* $z = 0.083$ *(or 341 Mpc), it is a nearby example of an Ultra-Luminous Infra-Red Galaxy or ULIRG.* Right*: Observer-frame spectrum of the best fit model to IRAS 14348-1447 at increasing redshifts, for a cosmology where* $H_\circ = 50\,km\,s^{-1}\,Mpc^{-1}$, $\Omega_\circ = 1$ *and* $\Omega_\Lambda = 0$*. Notice that at* $z = 0.08$ *by far most of the light coming from the galaxy is emitted in the far-infrared, i.e., it is re-emission of stellar light by dust grains. From:* Left*: HST/ADS image.* Right*: Devriendt et al. 1999, A&A 350, 381.*

11.1 Observations

It is not possible (yet) to bring representative samples of interstellar dust (see below) into the laboratory, and we must thus rely on remote observations. Our strongest constraints on interstellar dust come from observations of its interaction with electromagnetic radiation (see also Draine 2011):

- Wavelength-dependent attenuation of starlight by absorption and scattering, observable at wavelengths from 0.1 μm to 20 μm. This continuum extinction is strongest in the ultraviolet, and weakens towards the infrared, and includes a number of spectral features that provide clues to the grain composition (see Fig. 11.2 and Sect. 12.2).
- Polarization-dependent attenuation of starlight, resulting in a wavelength-dependent polarization of light reaching us from reddened stars.
- Scattered light in reflection nebulae. The dust in these clouds reflects the light of a nearby star or stars. These nearby stars are not hot enough to ionized the gas, such as in H II nebulae (also called emission nebulae; see Fig. 11.3).
- Strong attenuation of background starlight in interstellar clouds, causing such a cloud to appear as a dark patch in the sky relative to its surroundings. A famous example of a *dark cloud* is the Horsehead Nebula (see Fig. 11.3).

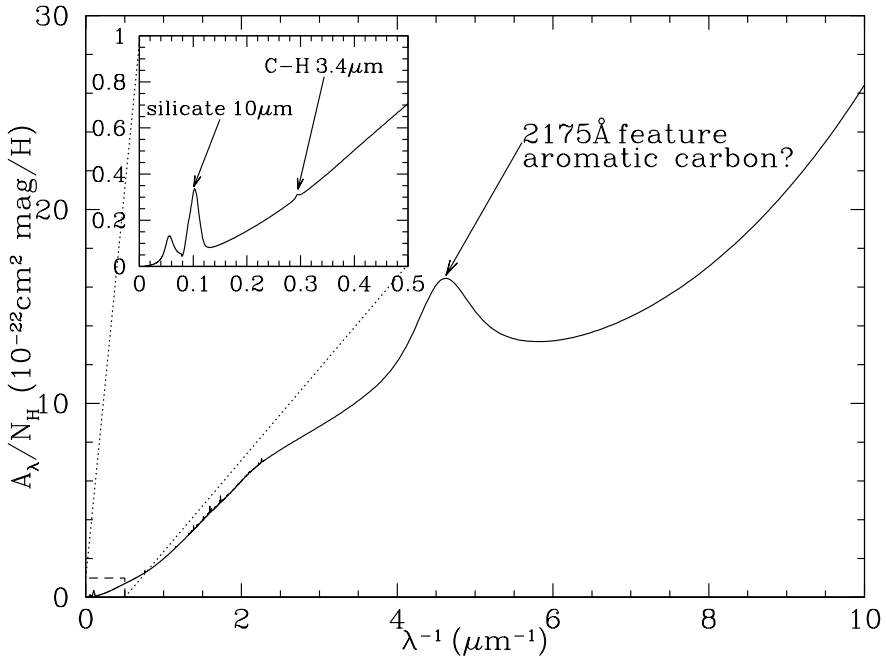


Figure 11.2: *Extinction versus inverse wavelength λ^{-1} on a typical sightline in the local diffuse ISM. The inset shows the extinction at $\lambda > 2\mu$m. Note distinct spectral features at 2175 Å, 3.4 μm and 10 μm that provide clues to the grain composition. Because atomic hydrogen absorbs strongly at $\lambda < 912$ Å, the extinction curve is shown up to about $\lambda^{-1} = 10\,\mu\text{m}^{-1}$. The dust responsible for interstellar extinction appears to be relatively well-mixed with the gas. For $R_V \sim 3.1$, $A_V/N_\text{H} = 5.3 \times 10^{-22}$ mag cm^{-2} H^{-1}, assuming all the gas is neutral.*

- ☐ Thermal emission from dust, at wavelengths ranging from about 2 μm to the sub-mm.
- ☐ Small-angle scattering of X-rays, causing 'scattered halos' around X-ray point sources.
- ☐ Microwave emission from dust, probably from rapidly spinning ultra-small grains.
- ☐ Luminescence when (a so far unidentified component of) interstellar dust is illuminated by ultraviolet starlight with efficient luminescence in the $500 - 1000$ nm spectral range – the so-called *extended red emission*.

In addition to this electromagnetic evidence, dust reveals itself also in other, less direct ways (see Draine 2011):

- ☐ Pre-solar grains preserved in meteorites – a selective and not-well understood sampling of the interstellar grains that were present in the solar nebula 4.65 Gyr ago. See Fig. 11.4.

Figure 11.3: *The Horsehead Nebula (or Barnard 33) in the emission nebula IC 434 is a dark nebula in the much larger Orion Molecular Cloud complex at a distance of 1500 lyr. The red or pinkish glow surrounding the horse head like shaped nebula is from an H II region, ionized by the nearby bright binary σ Orionis A (O9V) & B (B0.5V).*

- Depletion of certain elements from the interstellar gas, with the missing atoms presumed to be contained in dust grains (see Fig. 1.5).
- The observed abundance of H_2 in the ISM, which can only be understood if catalysis on dust grains is the dominant formation avenue (see Sect. 10.2).
- The temperature of interstellar diffuse H I and H_2, in part a result of heating by photo-electrons ejected from interstellar grains.

11.2 Lattice structure and chemical composition of dust

Crystalline and Amorphous lattice structure

We can distinguish two lattice structures of solids in space. Crystalline materials show a regular (ordered) structure both on short and long length scales. Amorphous materials also show some regularity, but distances to atoms and binding angles between atoms can vary. Glass is amorphous, so has short-range order but long-range dis-order. Lets take this material

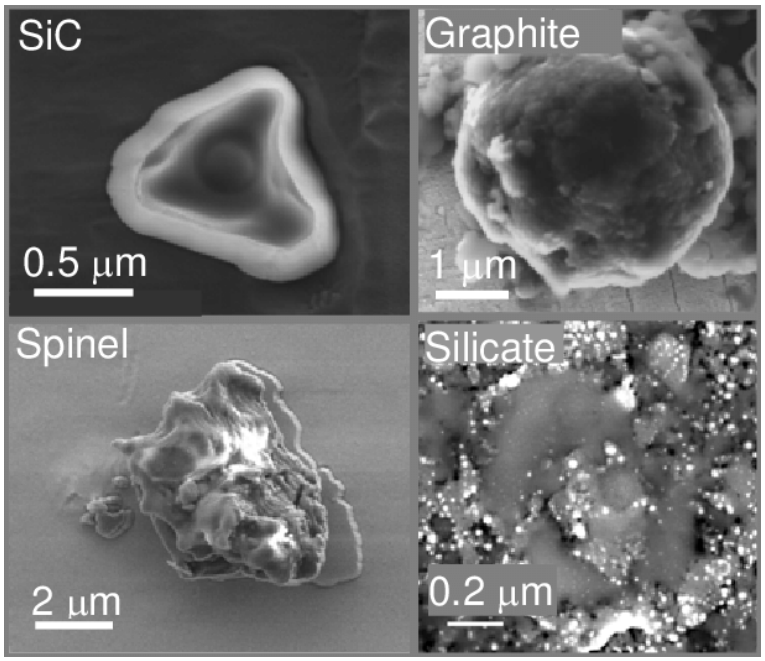


Figure 11.4: *Fotos of presolar grains found in primitive meteorites. Upper left: Silicon carbide (SiC). The isotopic signatures of these SiC inclusions point to an origin in a Type II supernova (SN) explosion. Upper right: Graphite from an asymptotic giant branch (AGB) star or a Type II SN. Lower left: Spinel ($MgAl_2O_4$) from an AGB star. Lower right: a grain consisting of an Al-rich core surrounded by a silicate mantle, also from an AGB star. From: Hoppe 2010. Photo credit: Max Planck Institute for Chemistry.*

as an example. The basic 'building block' or *network former* of glass is silica, SiO_2. Pure silica consists of a 3D network of tetrahedra in which Si is in the centre of the tetrahedron and every corner oxygen atom is shared with the adjacent tetrahedron (see Fig. 11.5). In a crystalline form, the tetrahedrons are in a regular geometry, while in an amorphous lattice regularity is lost at a length scale not much larger than that of the tetrahedron. In figure 11.6 this is shown schematically.

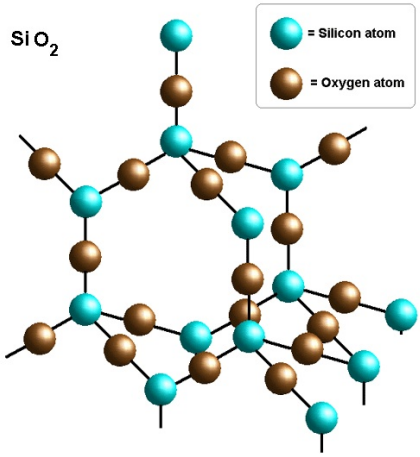


Figure 11.5: *The SiO_2 structure is composed out of Si-O_4 tetrahedrons which are connected at the corner O atoms under specific Si-O-Si bonding angles.*

So, what determines whether the lattice that forms is amorphous or crystalline? Figure 11.7 illustrates the volume-temperature phase diagram. A condensation point T_b exists at which a gas moves from the gas phase to the liquid phase (the subscript b denotes the reverse process, boiling). Contrary to the gas phase, in a liquid the molecules are locally bound: continuously bonds are created and broken. When

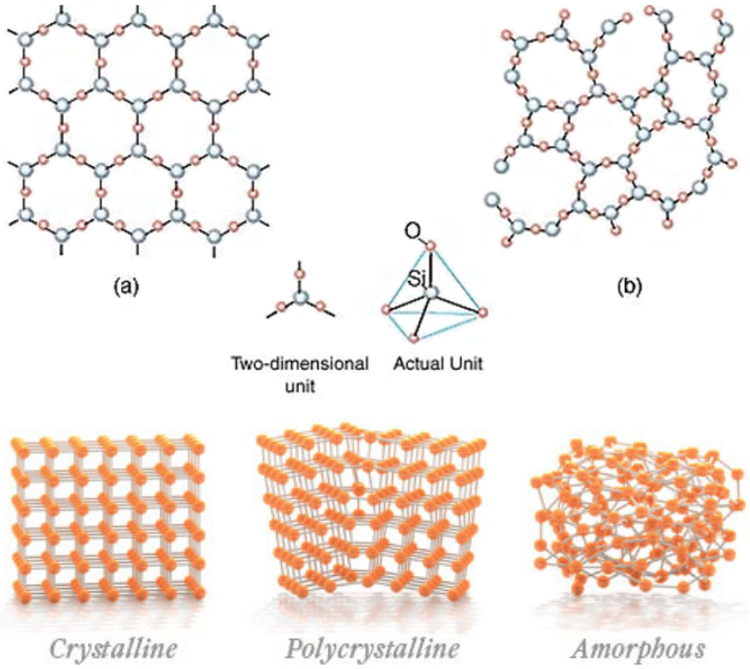


Figure 11.6: *Top: the atomic ordering in a crystalline (a) and amorphous (b) solid. The dots indicate the equilibrium positions of the atoms around which they can vibrate. The lines indicate the chemical bonds. Every silicon atom is in the core of an oxygen tetrahedron. In a crystalline solid, the particles are ordered in well-defined arrangements and create flat surfaces with definite angles to give highly regular shapes. In an amorphous solid, particles have no orderly structure which results in no well defined faces and shapes. Bottom: lattice structures that are intermediate between crystalline and amorphous are called polycrystalline. Though not typical for silicates, carbon grains often have a polycrystalline structure.*

the liquid cools *slowly* it passes the phase transition from liquid to solid at the freezing temperature $T_{\rm f} = T_{\rm m}$ (where the subscript m denotes the reverse process, melting). The atoms and molecules have enough time to find the energetically most favorable, crystalline, lattice structure. This process releases energy, according to the second law of thermodynamics. If the cooling is *fast(er)*, the material will remain in the liquid phase for a period of time before gradually solidifying in an amorphous material. Depending on the speed of cooling several amorphous structures can be formed (glass A and glass B in Fig. 11.7) If the temperature of the medium in which a solid forms is below the annealing temperature (somewhere in between $T_{\rm K}$ and $T_{\rm f}^{\rm B}$ in Fig. 11.7) the material can only solidify with an amorphous lattice form.

When gas pressure and temperature are very low, the liquid phase is skipped and a gas will go directly to the solid phase upon cooling (deposition). At typical interstellar matter conditions this is the most likely sequence. Moreover, as in the ISM temperatures are typically below $T_{\rm a}$ the forming solid will be amorphous.

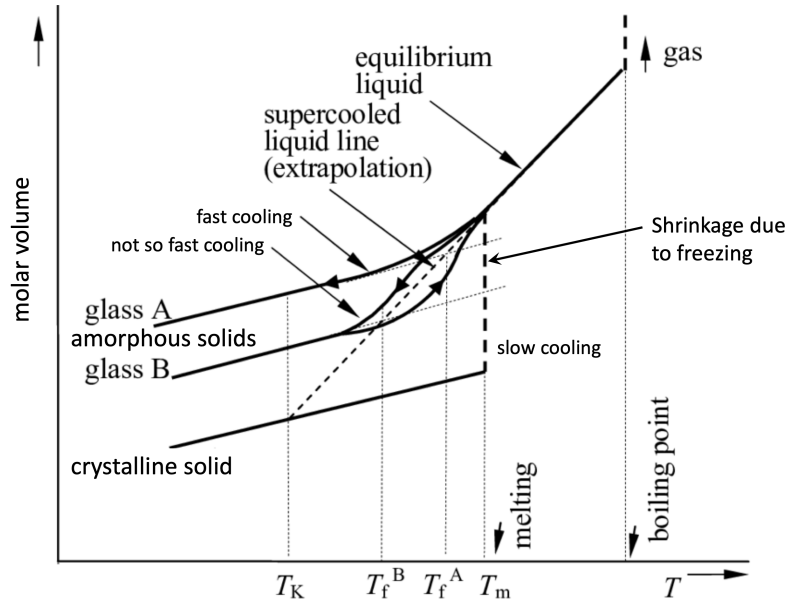


Figure 11.7: *Volume-temperature phase diagram. There are two routes along which a gas can solidify. Route 1 results in crystalline materials, route 2 is the fast cooling path and results in amorphous materials.*

The transition from an amorphous solid to a crystalline lattice structure is only possible by heating the material to a temperature above $T_{\rm a}$. This heating adds energy and increases the mobility of the atoms, resulting in an ordering of the lattice. This process is referred to as *annealing* and the temperature at which annealing can start ($T_{\rm a}$) is referred to as the glass temperature. An amorphous material that exhibits a glass transition may classify as a *glass*. In the solid phase a glass is usually transparent or translucent (i.e., semi-transparant). Glasses are typically brittle and consist of a mixture of silicates, including SiO_2.

An amorphous or crystalline lattice structure can be 'poluted' by the inclusion of other oxides (see Fig. 11.8). These oxides become part of the network and act as a stabilizer. There are two types of such 'pollutants'. *Intermediates* can replace the network formers. Examples of such intermediates are TiO_2 (see left panel Fig. 11.8) and Al_2O_3. *Modifiers* modify the network structure. Modifiers are usually present as positive ions (cations). Their positive charge is compensated by nearby non-bridging oxygen atoms, bound by one covalent bond to the network and holding one negative charge to compensate the positive ion nearby. In the right panel of Fig. 11.8 sodium is such a modifier. Some elements can act both as intermediates and modifiers. The presence of non-bridging oxygens lowers the relative number of strong bonds in the material and disrupts the network, lowering the freezing temperature.

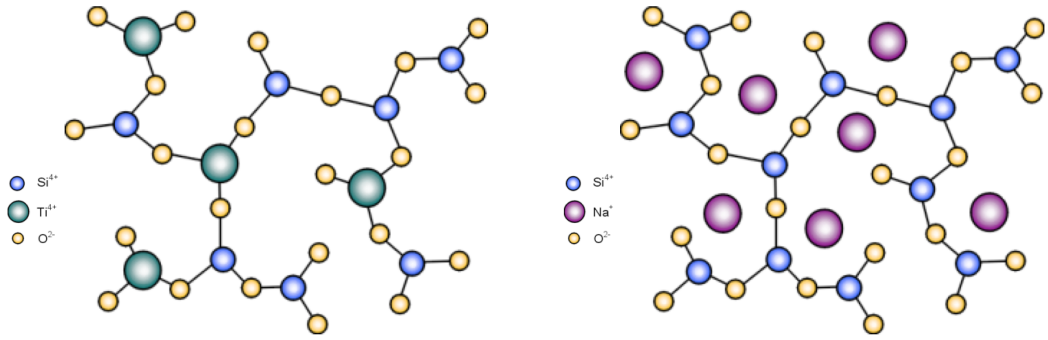


Figure 11.8: *Glassy silica (SiO_2; where Si is blue and O is yellow). Left: Glass with added titanium (green) as a network intermediate. Right: Glass with added sodium (purple) as a network modifier.*

11.3 Composition of interstellar dust

The available evidence indicates that the overall abundances in the ISM of our galaxy are close to the values in the solar photosphere. Though most abundant in number, hydrogen does not contribute appreciably to the grain mass. The way H is incorporated into grains is typically through C−H bonds, immediately implying that its contribution to the mass of the particle can only be at most $1/(12+1) \sim 10\%$. The nobel gasses helium and neon are chemically inert. Nitrogen too is a rather inert element when it comes to dust formation, and only seems to appear as an impurity in water (dominated) ice. Dust in the diffuse ISM appears to be ice free, but ices contribute a significant fraction of the dust mass in dark clouds. Shielding of the dust from the interstellar radiation field is probably needed to suppress H_2O removal by *photo-desorption* (the process in which atomic or molecular species leave the surface of a solid when it is exposed to light). Though H_2O is the dominant 'ice' species, ammonia (NH_3) is one of the secondary constituents.

Without the species mentioned, it must be that dust grains are build out of the most abundant condensible elements that remain: C, O, Mg, Si, and Fe. Figure 1.5 indeed shows that these elements are in fact under-abundant in the gas (depleted), with about 2/3 of the C and 90 percent or more of Mg, Si, and Fe presumed to be incorporated in dust grains in the typical diffuse interstellar cloud. In the figure, gas-phase abundances, relative to solar abundances, are plotted against the *condensation temperature* $T_{\rm cond}$, the temperature at which 50 percent of the element in question would be incorporated into solid material in a gas of solar abundances, at LTE at a pressure of $p = 10^3$ den cm^{-2} (Lodders 2003). $T_{\rm cond}$ indicates whether an element is able to form stable solid compounds in gas of solar composition. We see that there is a strong tendency for elements with high $T_{\rm cond}$ to be under-abundant in the gas phase, i.e. to have been incorporated in grain material.

With the elements providing the bulk of the grain population identified, the main candidates for dust material are (see Draine 2011):

- ☐ Silicates, i.e. *pyroxenes* (of composition $Mg_xFe_{1-x}SiO_3$) or *olivines* ($Mg_2Fe_{2-2x}SiO_4$ $(0 \leq x \leq 1)$.
- ☐ Oxides of silicon, magnesium, and iron (e.g. SiO_2, MgO, Fe_3O_4)
- ☐ Carbon solids (graphite, amorphous carbon, and diamond)
- ☐ Hydrocarbons (e.g. polycyclic aromatic hydrocarbons)
- ☐ Metallic iron

Other elements (e.g. Al, Ti, Cr) are also present in interstellar grains, but, because of their low abundances, they contribute only a minor fraction of the grain mass.

11.4 Observed spectral features of dust

Spectroscopy is a powerful way to measure the chemical composition of interstellar dust, as well as to derive information about grain size and shape. A good example is the 217.5 nm UV bump which is usually attributed to an electronic transition in the carbon lattice of a graphitic particle. At infrared wavelengths there are many spectral bands of dust species, that correspond to vibrational modes of the molecules that make up the grain (see Sect. 4.6). It is not simple to identify the precise chemical composition of the grains from the observed bands, because we can only compare the astronomical spectra to laboratory samples of materials present on earth. These materials may differ from those in space. However, in recent years a lot of effort has been invested in devising more realistic laboratory measurements, and this progress has resulted in the identification of many dust species. Depending on the temperature of the dust, we can observe the infrared vibrational resonances in absorption (cold dust in front of a bright background object) or emission (warm dust, usually circumstellar). In the case of Polycyclic Aromatic Hydrocarbons (PAHs) the bands are only seen in emission.

The 217.5 nm feature

The 217.5 nm (5.7 eV) bump is remarkably constant in wavelength, but its width can vary significantly between different lines of sight. The strength is well correlated with $E(B-V)$ (see Sect. 12.2). No other resonances in the UV are known. The feature is unpolarized, which means that its carrier is not part of the grain population causing extinction and polarization. The carrier of the 217.5 nm band must be abundant, in order to explain its strength. Graphite is the most likely candidate, but the problem is that a very specific grain size and shape is

Wavelength μm	Identification		Chemistry	Environment
Dust				
0.2175	Graphite		C	ISM
11.3	SiC		C	AGB
20.5	TiC clusters?		C	post-AGB shells
30	MgS		C	AGB — PN
9.7, 18	a-silicate		O	ISM, CS dust
13, 16.5	$MgAlO_4$	Spinel	O	AGB
11.3, 16.5 19.5, 23.5, 27.8, 33.5 69	c-Mg_2SiO4	Forsterite	O	AGB–PN, YSO
40.5	c-$MgSiO_3$	Enstatite	O	AGB, YSO
65	$CaSiO_7$	Diopside	O	post-AGB
110	Hydrous silicates		O	YSO
8.6, 20.6	SiO_2	Silica	O	AGB, YSO
Ices				
4.67	CO		C,O	MCs, PP disks
3, 43, 60	c-H_2O ice		O	AGB—PN, YSO
44	a-H_2O ice		O	MCs
2.70, 2.78, 4.27, 15.2	CO_2		O	MCs, protostars
4.92	OCS			Protostars
3.32, 7.67	CH_4			Protostars
2.27, 3.54, 3.85, 3.94, 4.1, 6.85, 8.9, 9.7	CH_3OH			Protostars

Table 11.1: Incomplete overview of identified dust bands in the mid- and far-infrared part of the spectrum. MC = molecular cloud; CS = circum-stellar; PP = proto-planetary; PN = planetary nebula; a- = amorphous; c- = crystalline.

required to match the observations, in particular the observed and very constant wavelength of the band. Its carrier therefore remains debated, and alternatives have been suggested, such as OH^- on small silicate grains, see Steel & Draine (1987) and Bradley et al. (2005); and a combination of organic carbon and amorphous silicates, see Bradley et al. (2005).

Silicate features at 9.7 μm and 18 μm

There is a conspicuous absorption feature at 9.7 μm that is detected throughout the galaxy and in other galaxies. Silicate minerals generally have strong absorption resonances due to the Si−O stretching mode near that wavelength, and it seems virtually certain silicates are the carrier of the feature. This conclusion is strengthened by the fact that the 9.7 μm feature is seen in the outflows from oxygen-rich stars (which would be expected to condense silicate dust) but not in the outflows from carbon-rich stars (expected to condense carbon based dust).

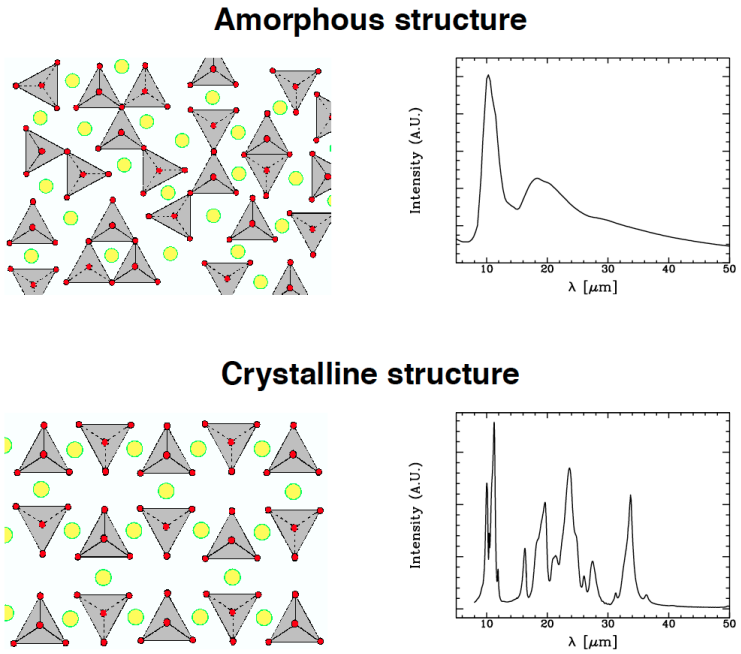


Figure 11.9: *A possible atomic structure of a disordered (or amorphous) silicate and that of an ordered (or crystalline) silicate together with their typical infrared emission spectra. The projected Si-O_4 tetrahedrons are shown in gray (oxygen atoms are in red, and the silicon atom is hidden from view by the center oxygen atom that is above the plane of the paper) and the yellow circles are the metal cations. Note the many sharp features in the crystalline silicate spectrum and the two broad bumps at about 10 μm and 19 μm for the amorphous silicate spectrum. From: Molster & Kemper (2005).*

The interstellar 9.7 μm feature is seen both in emission (e.g. in the Trapezium region of Orion; Gillett et al. 1975) and in extinction in the interstellar medium (Roche & Aitken 1984). Sightlines within a few kpc from the Sun have (Draine 2003)

$$A_V = (18.5 \pm 2) \times \tau_{9.7}, \tag{11.1}$$

where A_V is the visual extinction (see Eq. 12.11) and $\tau_{9.7}$ the optical depth at 9.7 μm. However, sightlines to sources near the Galactic Center have $A_V = (9 \pm 1) \times \tau_{9.7}$ (Roche & Aitken 1985). This correlation breaks down for lines of sight that pass through molecular clouds: with increasing column of dust along these lines of sight, the silicate band strength no longer increases in proportion. Clearly dust properties (size, chemical composition) change in molecular clouds.

Near 18 μm, interstellar dust shows another feature, attributed to the O$-$Si$-$O bending mode in amorphous silicates.

The SiO *molecule* has its fundamental vibrational stretch resonance $\nu_{\rm osc}$ at 8.13 μm (see Eq. 4.28). The band shape and central wavelength of the resonance in the solid (at 9.7 μm)

are different from that of the gas-phase molecule because the forces acting on the individual atoms are different due to nearest neighbor effects, shifting the band to longer wavelengths and broadening it. In addition, rotational motions are not possible in the solid lattice. As a result of these differences, amorphous materials only show the most fundamental resonances in their spectra. An overview of silicates is given in Table 11.1. Common cations incorporated in the lattice structure of silica tetrahedrons are Mg, Fe, Al, and Ca. Important groups of silicates are the *olivine group*, with chemical formula (Mg, Fe, ..)$_2$SiO$_4$, and the pyroxene group, with chemical formula (Mg, Fe, ..)SiO$_3$.

Silicate crystals

At first it was generally assumed that all cosmic silicates were of amorphous structure. Thanks to the *Infrared Space Observatory* (ISO), launched in 1995, we now know that crystalline silicates are ubiquitous in the Galaxy – though they appear absent in the diffuse galactic ISM (a stringent limit of $< 1\,\%$ of the total silicate mass has been placed on it; Kemper et al. 2004). They do appear to be present in the ISM of distant starburst galaxies. The resonances in crystalline silicates are sharply peaked relative to those of amorphous silicates. Crystalline silicates can also be distinguished from amorphous silicates due to the presence of lattice modes at wavelengts $\gtrsim 25\,\mu$m (see Fig. 11.9). Table 11.1 includes some crystalline silicates.

A nice example of crystalline silicates can be found in Fig. 11.10, which shows the ISO spectrum of the planetary nebula NGC 6302. This object has recently evolved off the Asymptotic Giant Branch (AGB) into the Planetary Nebula phase. The old AGB wind is still visible in the spectrum as the cold dust continuum with many resonances of crystalline silicate dust superimposed. Crystalline silicates are also found in the planet-forming disks around young stars and in Solar System comets.

Ices

In dark molecular clouds and near proto-stars, a number of additional absorption features appear, most notably a strong band at 3.1 μm which is produced by the O−H stretching mode in H_2O ice. Other common ices are CO and CO_2 ice. These features are not seen on lines of sight that pass only through diffuse interstellar clouds, even when the extinction is large. So, dust in the diffuse ISM appears to be ice-free. It indicates that ice is present only in regions that are shielded from the diffuse starlight background. The dust shielding is probably needed to suppress H_2O removal by photo-desorption. Lines of sight towards young stars often show a mix of silicate and ice absorption bands (see Fig. 11.11).

The 3.4 μm feature

There is a broad absorption feature at 3.4 μm (see Fig. 11.2) that is almost certainly due to the C−H stretching mode of hydrocarbons (see Sect. 4.6 and Fig. 11.12). Hydrocarbons with

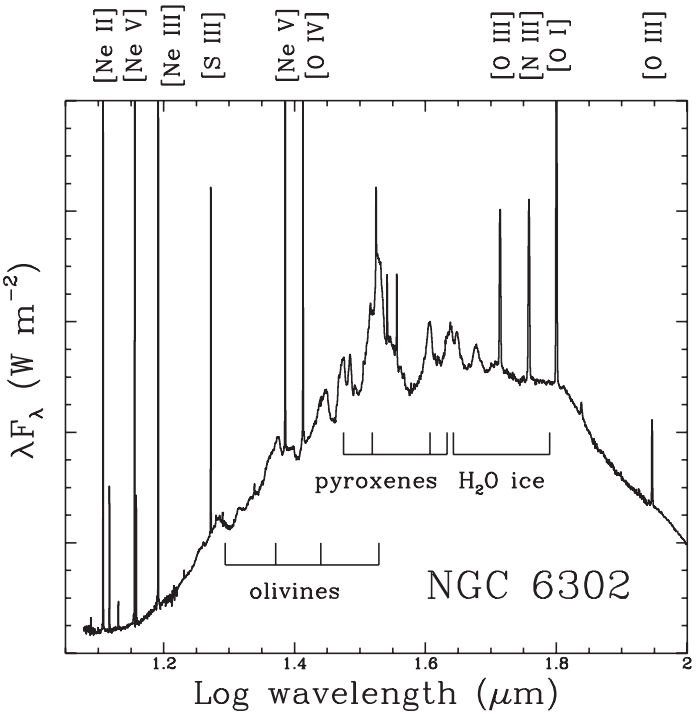


Figure 11.10: *The ISO spectrum of the planetary nebula NGC 6302, showing strong emission bands of crystalline silicates and of crystalline water ice. In addition, the spectrum also shows very strong forbidden fine-structure emission lines of the ionized part of the nebula.*

a mixed aromatic (ring) and aliphatic (chain) character provide a good fit to the observed feature, but the exact aromatic / aliphatic ratio remains uncertain. Somewhat surprisingly, the 3.4 μm C−H feature is found to be weaker (relative to the overall extinction) in dark clouds than in diffuse clouds (Shenoy et al. 2003), which has been interpreted as evidence that the C−H bonds responsible for the 3.4 μm feature are destroyed in molecular clouds, perhaps as a result of cosmic ray irradiation, and regenerated when carbonaceous grains are exposed to atomic hydrogen in diffuse clouds (Mennella et al. 2003).

PAH features

A number of emission bands are very frequently observed in many different environments, ranging from the diffuse ISM to H II regions, reflection nebulae, C-rich planetary nebulae (see Fig. 11.13), photon dominated regions, planet-forming disks around Herbig AeBe stars and T Tauri stars, and even in the integrated light of entire galaxies (see Fig. 1.2). The strongest bands are found at wavelengths of 3.3, 6.2, 7.7, 8.6, 11.3 and 12.7 μm (see Figs. 1.3 and 11.12). A larger number of weaker bands are also found. This family of emission bands has been known since the 1970s. Laboratory studies and detailed observations with ISO and Spitzer

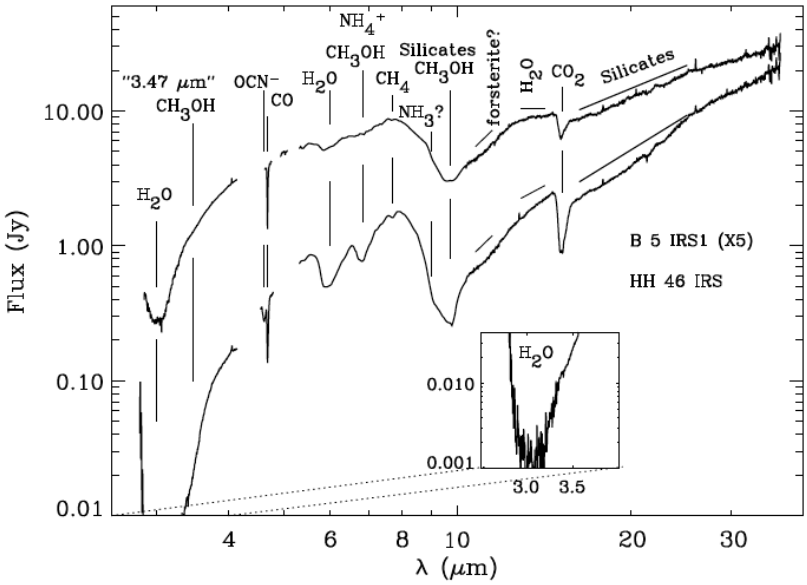


Figure 11.11: *Spitzer Space Telescope spectra of young low-mass stars embedded in molecular cloud material with substantial amounts of ices, seen as absorption bands in the spectrum. From: Boogert et al. 2004, ApJS 154, 359.*

Space Telescope have convincingly shown that the carriers of these bands are most likely *Polycyclic Aromatic Hydrocarbons* (PAHs). The observed bands coincide with strong bending and stretching resonances of these PAH molecules, see Fig. 11.12.

The PAH molecules have as basic 'building block' a benzene ring (C_6H_6). Benzene has as additional property over simpler hydrocarbons, such as methane (CH_4, that its ring-like structure allows the electrons of the π-orbitals to delocalise over the entire carbon ring. Molecules with this property are called *aromatic*. *Polymerisation* of two or more benzene rings creates 2- and even 3-dimensional structures. The edges of the structures can have H-atoms attached (like in benzene). Such molecular structures can show different kinds of vibrational resonances, such as C-H (in-plane stretch, out-of-plane stretch, in-plane bending) and C-C (stretch). The family of PAHs is large and ranges from its smallest member naphthalene ($C_{10}H_8$) to very large molecules such as $C_{384}H_{48}$.

The precise nature of the PAH molecule is not crucial in determining the wavelength of the resonance, so that it is still not clear what the detailed structure of the dominant PAH molecules in different environments is. Observations suggest that typical PAH molecules have about 60 C atoms, and a varying amount of C-H bonds at the edges of the molecules.

Interestingly observations of PAH molecules show that the *relative* strength of the PAH emission bands in e.g. reflection nebulae hardly depends on the distance to the star that excites the molecules. For dust in thermal equilibrium a sharper drop in band strength would be ex-

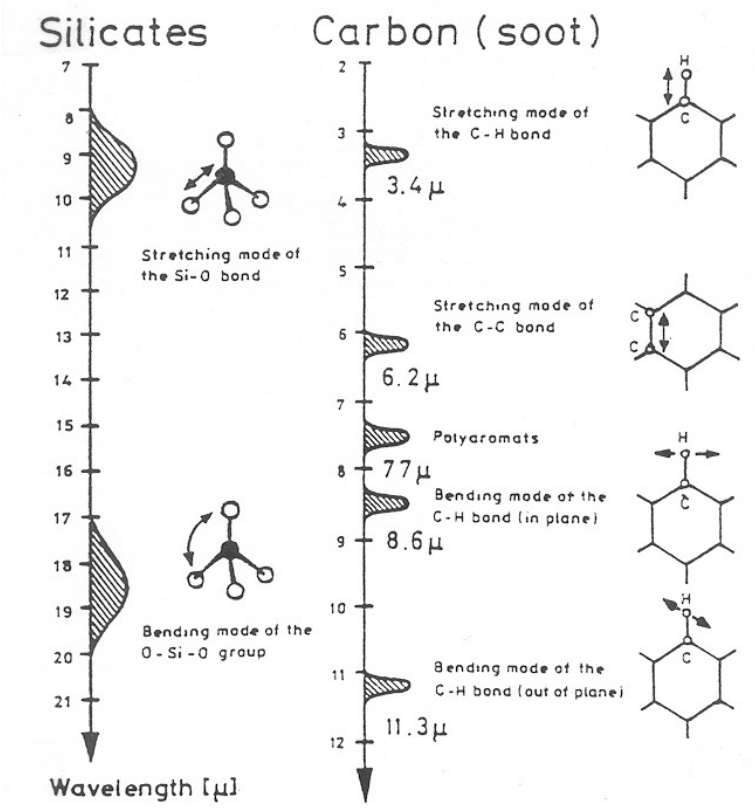


Figure 11.12: *Left: the stretch resonance of Si-O together with the bending resonance of the O-Si-O bond in a tetrahedron give rise to bands at 9.7 and 18 μm that can be in emission or absorption. Right: the most important resonances in Polycyclic Aromatic Hydrocarbons, PAHs.*

pected because the temperature of the grains drops with distance to the star heating the grains. However this effect is not observed. The PAH bands to first order only weaken due to the dilution of the radiation field as the distance to the central star increases. Even in very cold dust regions PAH emission is readily found. Obviously, we are not dealing with the usual resonances in dust grains that are in thermal balance, but rather with fluorescence of molecules that are excited by UV and optical photons. These short wavelength photons are absorbed by electrons in the molecule, and internal energy conversion within the molecule results in the emission of infrared photons in the vibrational resonances of the molecule. This fluorescence mechanism explains why the PAH bands are never seen in absorption, only in emission. The molecules de-excite on very short timescales and usually the UV radiation field is so dilute that the time to the next photon absorption is hours to days (see Sect. 12.4). It means that

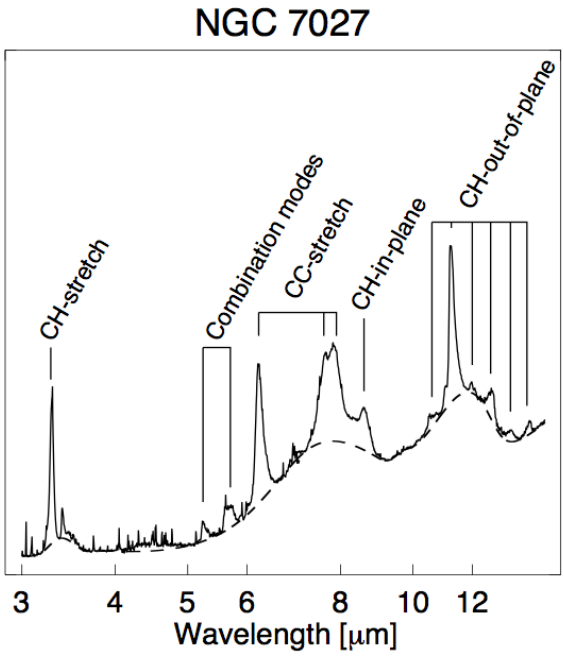


Figure 11.13: *The mid-infrared spectrum of the carbon-rich planetary nebula NGC 7027. The emission bands in the spectrum are due to PAH molecules in the ejecta of the star that are illuminated by UV photons of the hot white dwarf.*

PAHs are transiently heated to a few 100 K and quickly cool to very low temperatures. Only in very intense radiation fields the cooling timescale becomes similar to the photon absorption timescale. Under these conditions the short-wavelength bands will emit stronger, simply because the molecule is on average hotter. Other effects that influence the relative strength of the PAH bands are the size of the molecule, ionization properties and dehydrogenation.

Graphite and diamond

As pointed out above, graphite seems to be the most likely candidate of the pronounced 217.5 μm band. Indeed, the observed profile of the feature is well fitted by theoretically calculated extinction cross sections for either 200 Å graphite spheres or for graphite prolate (i.e. flattened) spheroids with a size of 30 Å and an axial ratio of 1.6. The sensitivity of the absorption profile to the detailed characteristics of the grain poses a problem for all of these fits in view of the observed constancy of the peak position (but variable width) of the 217.5 μm feature. It should also be mentioned that not only graphite may fit the feature. Good fits can also be obtained with laboratory measured extinction spectra of $\sim$100 Å hydrogenated amorphous carbon grains – containing 33 % H atoms by number – which share the aromatic character of their bonding with graphite.

Diamond nano-particles are relatively abundant in primitive Solar System meteorites. Based

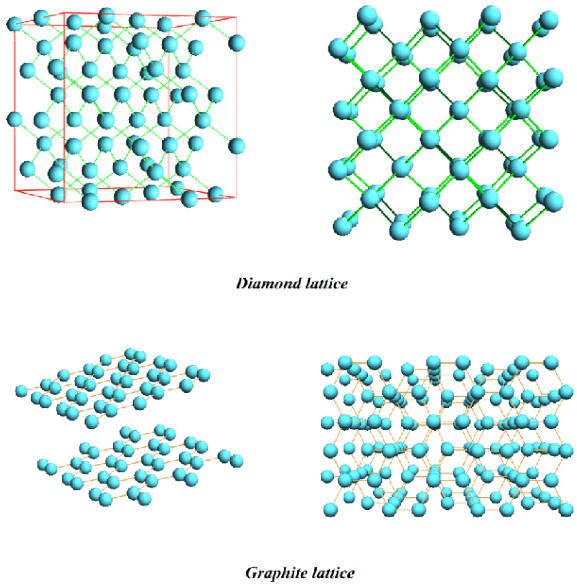


Figure 11.14: *Graphite consists of parallel sheets of so-called sp^2-bonded carbon. A single (sheet) of carbon atoms is known as* graphene*; each carbon atom in graphene has three nearest neighbors (one of which is a double bond, i.e. the bond has a aromatic character). The graphite sheets are weakly bonded to one another by van der Waals forces. Diamond consists of sp^3 bonded carbon atoms, with each carbon atom is bonded to four equidistant nearest neighbors.*

on isotopic anomalies associated with them, it seems likely that some fraction of the nano-particles is of pre-solar origin, and thus was present in the ISM prior to the formation of the Sun. Therefore, *some* nano-diamond is presumable present in the ISM today, but its abundance is not known (see Jones & D'Hendecourt 2004).

Diffuse Interstellar Bands

The spectrum of the interstellar medium shows weak absorption bands from near-UV to the near-IR (see Fig. 11.15). The bands have widths from 0.5–30 Å, which is significantly broader than the (Doppler) width of atoms, ions, or small molecules. These bands were discovered a century ago (Heger 1922) and have been dubbed the *diffuse interstellar bands*. Their interstellar nature was established by Merrill in 1934 (Merrill 1934). The strongest DIB falls at 4430 Å. The DIB spectrum shows strong variations in relative strength, as illustrated for the well-known 5780 Å and 5797 Å DIBs in Fig. 11.15. Hobbs et al. (2009) report a total of 414 DIBs between 3900 Å and 8100 Å.

It seems likely that at least a substantial part of the DIBs may be due to free-flying large molecules, possible ionised. Support for this hypothesis comes from laboratory experiments likely identifying C_{60}^+ as the carrier of the 9577 Å and 9632 Å DIBs (Campbell et al. 2015). Further support comes from high resolution spectra of the 5797 Å feature showing intrinsic

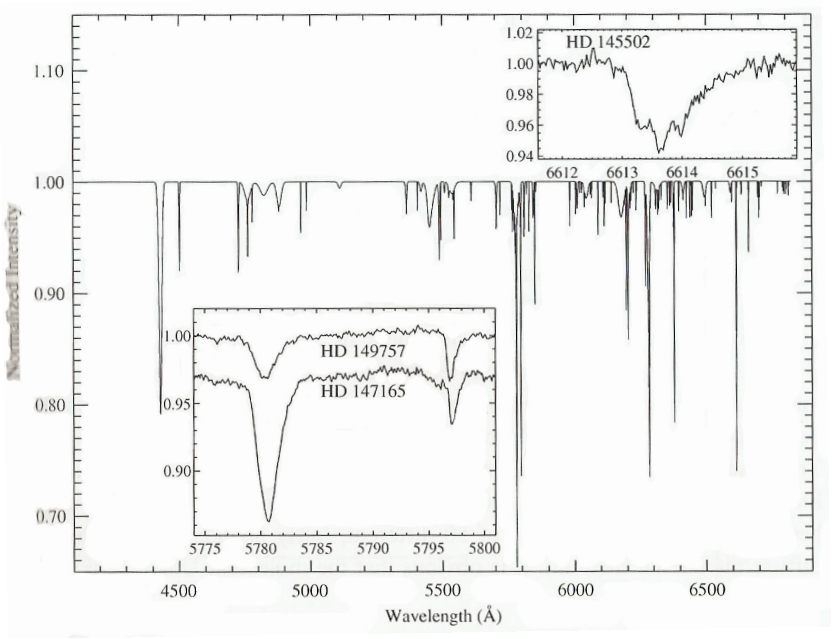


Figure 11.15: *The synthesized 4300-6800 Å spectrum of the diffuse interstellar bands as derived from observations of the star BD+63 1964 (Ehrenfreund et al. 1997) illustrates the great variety in relative strength and width of these absorption features. Note the strongest of all DIBs at 4430 Å. The top insert shows the detailed profile of the 6614 Å DIB and its associated substructure observed towards the star HD 145502. The bottom insert illustrates the large variations in the strength of the DIB bands relative to each other for the two well-studied DIBs at 5780 Å and 5797 Å on the basis of observations of the stars HD 149757 and HD 147165. Note the strongest of all DIBs at 4430 Å. From: Tielens (2005), courtesy of Jan Cami.*

ultra-fine structure. The DIBs may represent absorption in an electronic transition of such molecules, and the sub-structure may reflect rotational substructure. If so, one would expect that a given molecule would have multiple absorption lines due to different vibrational states of the electronic excited state. McCall et al. (2005) found what appears to be a nearly perfect correlation between the strengths of DIBs at 6196.0 Å and 6613.6 Å, suggesting that these may be two absorption features produced by a single absorber – similar as the two DIBs associated with C_{60}^{+}. But such correlations in DIB behaviour are exceptions. The observation that generally the strong bands vary in relative strength from one sightline to another implies that each of them represents a single species and hence that there are many different (more than 50 or so), fairly abundant molecules in the diffuse ISM.

Clearly, the DIB carriers have to be readily made or survive easily in the harsh environment of the diffuse ISM. Regarding the first option, the production mechanism of DIBs may be linked to the reservoir of dust particles in the sense that UV illumination of grains may slowly evaporate them, releasing (among other material) DIB carriers. Regarding the latter, small molecules (with few atoms only) are usually efficiently destroyed by the interstellar UV radi-

ation field, however (some) more complex molecules may be more refractory to UV illumination. They might be produced locally or carried over from their place of origin, implying such DIBs are not directly linked to the local dust reservoir.

Exercise 11.1

This could be a nice exam question. The $3\,\mu$m water ice absorption feature is detected in the direction of sources *behind* the Taurus region if the visual extinction $A_V > 3$ mag, which is called the threshold value for ice formation. The observations indicate the presence of mantles on ice grains which mainly consist of amorphous water ice. In this exercise we compute how these mantles have formed from molecules in the gas phase.

a) The accretion rate of atoms and molecules on grains depends upon the relative velocity between grains and gas atoms. The velocity distribution of particles is given by the Maxwell distribution. The distribution of the absolute velocity $v = |\mathbf{v}|$ of particles with mass m is given by equation (2.61). Derive the mean relative velocity between the dust grains and gas particles.

Remember that the two particle problem becomes a one particle problem by replacing the masses m_1 and m_2 by the reduced mass $1/m_r = 1/m_1 + 1/m_2$. Also:

$$\int_0^\infty e^{-ax^2} x^3 \mathrm{d}x = \frac{1}{2} a^{-2} \tag{11.2}$$

b) Derive an expression for the rate at which the mass of a grain grows as a consequence of the accretion of atoms and molecules. Make use of the mean velocity derived in the previous item, and use the following quantities: the grain radius a; the number of atoms and molecules per unit volume n; the sticking probability S; the gas temperature T, and the mean mass of the gas particles m.

c) Derive an expression for the speed at which the thickness of the mantle grows, da/dt. The specific density of the ice is $\rho_{\rm ice}$. Make use of your answer in (b). Do grains which were initially larger also grow a thicker mantle?

d) The grain mantles in the Taurus region have a thickness of $\Delta a \simeq 0.02\mu$m and a density $\rho_{\rm ice} = 0.75\,\mathrm{g/cm^3}$. Each particle that collides with the grain sticks ($S = 1$). Use Table 11.2, the result of (c) and assume $n(\mathrm{H_2}) = 2 \cdot 10^3\,\mathrm{cm^{-3}}$ and $T_{\rm gas} \simeq 50\,\mathrm{K}$ to show that this result can only be correct if we assume that the H_2O molecules are formed on the surface by surface reactions between O and H atoms, and not by direct condensation of H_2O molecules from the gas phase.

Table 11.2: *Abundances relative to $n(\mathrm{H_2})/2$ in a molecular cloud after 10^5 years. From: Hasegawa & Herbst (1993).*

Species	Abundance	Species	Abundance	Species	Abundance	Species	Abundance
H	6.0×10^{-1}	He	2.8×10^{-1}	C	6.1×10^{-5}	N	3.1×10^{-5}
O	2.2×10^{-4}	OH	1.1×10^{-8}	H_2	1	CO	4.3×10^{-5}
H_2O	5.8×10^{-8}	SiO	1.7×10^{-10}	NH	5.0×10^{-9}	CH	1.3×10^{-8}

12

Physics of interstellar dust

Dust grains can both absorb and scatter electromagnetic radiation. The effect of absorption and scattering together is called *extinction* (see Sect. 2.3). Light which passes through interstellar space suffers from interstellar extinction due to its interaction with the dust that is present. When the amount of extinction is small, the amount of dust along the line of sight (its column density) can easily be derived from observations. However when extinction is high, the amount of light which is 'lost' along the line of sight is very high and it may be difficult to constrain the column density. Examples of lines of sight with high extinction are the galactic centre that is obscured by 30 magnitudes of optical extinction, and lines of sight passing through dense molecular clouds. These regions with high extinction cannot be studied at optical wavelengths, even with the most powerful telescopes.

Dust particles play an important role in determining extinction due to their size. Electromagnetic radiation in general has a much longer wavelength than individual atoms or molecules, and light can pass without much difficulty. Only in spectral lines photons can be efficiently absorbed, and of course in H II regions Lyman continuum photons are strongly absorbed. Photons with wavelengths longward of the Lyman edge (at 912 Å) can efficiently be absorbed and scattered by dust particles of comparable size. Smaller dust particles are more abundant than large ones, and so extinction is larger at short (ultraviolet) wavelengths, and decreases towards longer (infrared) wavelengths. At infrared wavelengths the extinction to the galactic centre is only a few magnitudes, low enough to allow detailed studies of that important region in the Galaxy using imaging techniques and spectroscopy. Realize though that even at very long wavelengths (up to 40 μm) extinction can be significant – e.g. in star-forming regions.

In this chapter we discuss diagnostic methods to constrain the amount of dust in clouds or along line-of-sights toward stars, as well as some basic properties of the grains. Two sections are devoted to obtaining the temperature of dust grains.

12.1 Growth and destruction of grains in the ISM

Grain growth

Grains form in the outflowing gas from cool stars as well as in other environments, e.g. during supernova explosions. Do these grains grow significantly once deposited in the ISM? Consider a spherical dust grain at rest in the ISM with radius a that grows by the addition of species i (an atom or molecule) that has particle density n_i, mass m_i, and mean thermal speed v_i. The rate at which atoms stick to the surface is

$$R_{\rm stick} = 4\pi a^2\, v_i\, n_i\, S, \tag{12.1}$$

where S is a dimensionless *sticking probability* that describes the fraction of particles that hit the grain that actually stick to it. The mass m of the grain will grow at a rate

$$\frac{{\rm d}m}{{\rm d}t} = m_i\, R_{\rm stick} = m_i\, 4\pi a^2\, v_i\, n_i\, S, \tag{12.2}$$

If the grain has volumetric mass density $\rho_s = m/(4\pi\, a^3/3)$ then

$$\frac{{\rm d}a}{{\rm d}t} = \frac{S\, n_i\, m_i\, v_i}{\rho_s}. \tag{12.3}$$

We can easily integrate this equation if n_i and v_i do not depend on time. We get

$$a(t) = a_\circ + \frac{S\, n_i\, m_i\, v_i}{\rho_s} \cdot t, \tag{12.4}$$

where $a_\circ$ is the size of the nucleation seed, i.e. the size to which the grain could grow at the location where it was formed. For a typical grain in the ISM, $\rho_s = 1\,{\rm gr\,cm^{-3}}$. In the cool neutral medium $n_{\rm H} \sim 10\,{\rm cm^{-3}}$. If the species that sticks to the grain is 1000 times less abundant than hydrogen, $n_i = 0.1\,{\rm cm^{-1}}$. If it is 10 times as massive ($m_{\rm H} = 1.67 \times 10^{-24}\,{\rm gr\,cm^{-3}}$), it will have a mean thermal speed of $\sim$0.5 kms (see Eq. 2.62) for a typical temperature of the CNM. Assuming the nucleation seed to be negligibly small and $S \sim 1$, the time for the grain to grow to a size $a = 1\,\mu{\rm m}$ is $\tau_{\rm growth} = a/(da/dt) \sim 10^9\,{\rm yr}$. This is a long time compared with other relevant mechanisms and implies that grains do not grow significantly in the ISM. They first need to enter denser regions for growth to occur.

Grain destruction by evaporation and sputtering

If grains accrete icy mantles then these are fairly readily destroyed by *evaporation*, should the grain temperature rise when the grains pass near to a star. Refractory materials such as hydrocarbons, graphite, silicates, or silicon carbide are very durable and probably would survive such a star passage quite easily. These particles are however vulnerable to *sputtering* by

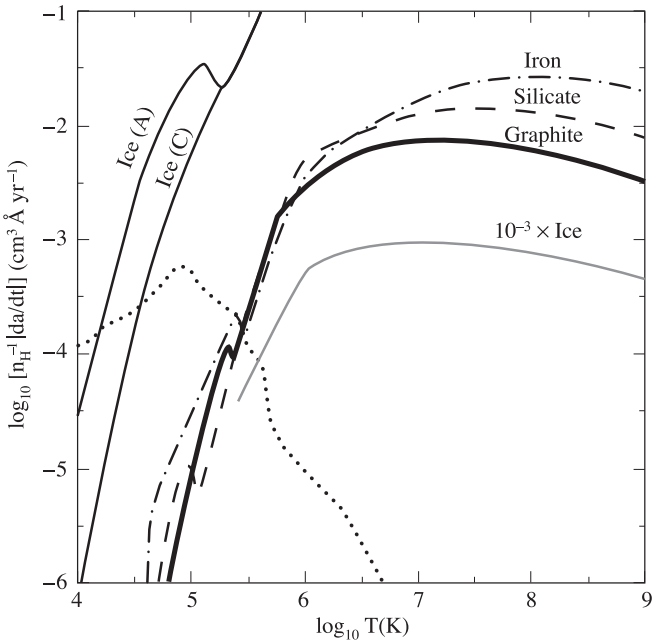


Figure 12.1: *Thermal sputtering rates (Draine & Salpeter 1979). For* $10^6 \lesssim T \lesssim 10^9$ *K all three refractory materials (iron, silicate, graphite) have* $1/n_{\rm H} \cdot |da/dt| \approx 10^{-2}$ *cm*^{3} *Å yr*$^{-1}$. *From: Draine (2011).*

high-speed atoms. Sputtering is a process in which an incident atom has sufficient momentum to knock one of the lattice atoms – or a small cluster of lattice atoms – completely out of the lattice. This typically needs velocities greater than about 50 $\mathrm{km\,s^{-1}}$ and these are probably attained in shocks such as caused by supernova explosions.

At an elementary level, we may regard sputtering as a simple momentum transfer proces, in which case the maximum energy transfer $\Delta E_{\rm kin}$ from the projectile p to the target t is

$$\Delta E_{\rm kin} = \frac{4 m_p\, m_t}{(m_p + m_t)^2}\, E_p = \eta\, E_p, \tag{12.5}$$

where m_p and m_t are the atomic masses of the particles involved (see exercise 12.2), and E_p is the projectile energy. If the binding energy of the atom close to the surface of the solid is $U_\circ$, then no sputtering can occur if $\Delta E_{\rm kin} < U_\circ$. In other words, there is a threshold energy below which sputtering can not take place.

To describe the outcome of laboratory experiments of sputtering one introduces the *sputtering yield* $Y_{\rm sput}$. Suppose we have a flux of projectile particles incident on a target surface, then the yield tells us how many atoms from the target are ejected per incident projectile. The yield actually depends on a number of factors, including the nature and momentum of the bombarding particles, the nature of the surface, and the angle at which the particles strike the

surface. It is complications such as these that make sputtering so difficult to describe. Still, to mention a number: $Y_{\rm sput}$ is $\sim$0.01 for 5 keV protons incident on graphite.

Figure 12.1 shows the outcome of laboratory experiments of thermal sputtering rates for ices and refractory grains. For $10^6 \lesssim T \lesssim 10^9$ K the calculated sputtering rates for graphite, silicate, or iron grains can be approximated by

$$\frac{{\rm d}a}{{\rm d}t} \approx -\frac{1\times10^{-6}}{1+(T/10^6\,K)^{-3}} \cdot \left(\frac{n_{\rm H}}{\rm cm^{-3}}\right) \quad \mu m\,{\rm yr}^{-1}, \tag{12.6}$$

to which one may associate a grain lifetime

$$\tau_{\rm sput} = \frac{a}{|da/dt|} \approx 1\times10^5\ \left[1+(T/10^6\,K)^{-3}\right] \left(\frac{a}{0.1\,\mu{\rm m}}\right) \cdot \left(\frac{\rm cm^{-3}}{n_{\rm H}}\right) \quad {\rm yr}. \tag{12.7}$$

In a supernova remnant with $n_{\rm H} = 1\,{\rm cm}^{-3}$ and $T = 10^6$ K, a grain with initial radius $a = 0.1\,\mu$m could survive for $\sim 10^5$ yr. This coincides with the typical lifetime of such a remnant, suggesting that though grains may form in the expanding remnant gas most of these grains likely will be destroyed before the remnant has dissipated. In the x-ray emitting intra-cluster gas of the Coma cluster, with $T \approx 10^8$ K ($kT \approx 8$ eV) and $n_{\rm H} \sim 0.003\,{\rm cm}^{-3}$, a 0.1 μm dust grain would have a lifetime $\tau_{\rm sput} \sim 3\times10^7$ yr. This is short compared to the age of galaxy clusters, hence these may be expected to be dust free.

A special type of sputtering is *chemical sputtering* in which sputtering takes place via a series of chemical reactions on the grain surface. For this proces the threshold energy is very low or none-existent. As an example we mention the surface of an ice grain that may erode by virtue of the bombardement of the grain by protons: $H_2O + H \rightarrow OH + H_2$.

Grain destruction by grain-grain collisions

Grains may collide with each other. The outcome of such a collision will depend on the circumstances: they may stick, shatter (each other), or destroy one or both particles. The latter is referred to as *evaporation*. Laboratory experiments and theory suggest that when particles shatter, the mass distribution of fragments $n(m)$ is given by

$$n(m)\,dm \propto m^{-\alpha}\,dm \tag{12.8}$$

where α is a constant. For asteroid belt particles, for instance, $\alpha \simeq 1.8$ for sizes $\sim$1000 km downwards. In the context of interstellar dust grains, however, it is more convenient to consider the *size* distribution, i.e. the number $n(a)\,da$ of grains with radius a in the range $[a, a+da]$. For a specific density of the grain material $\rho_{\rm s}$, we have that

$$dm = 4\pi\,a^2\,\rho_{\rm s}\,da. \tag{12.9}$$

We may now write

$$n(a) = n(m)\,\frac{{\rm d}m}{{\rm d}a} \propto m^{-\alpha}\,a^2 \propto a^{-3\alpha+2}. \tag{12.10}$$

A power-law distribution of grain masses also represents a power-law distribution of grain radii. Thus, in the asteroid belt the size (rather than mass) distribution has exponent $-3\alpha+2 = -3.4$. For interstellar grains, Mathis, Rumpl, and Nordsieck applied a power-law grain size model for graphite and silicate grains in order to fit the interstellar extinction curve in the 50−2500 Å range (see Fig. 12.2 for this curve). They find a power-law slope of -3.5, close to what is derived for the asteroid belt. This so-called MRN distribution is widely applied in studies of the ISM.

12.2 Interstellar extinction

It is custom to characterize the attenuating effects of dust by the *extinction* A_λ at wavelength λ, measured in magnitudes, defined by

$$A_\lambda = m - m_\circ = -2.5 \log \left[I_\lambda / I_\lambda^\circ \right], \tag{12.11}$$

where I_λ is the observed specific intensity, and I_λ° the specific intensity that would have been observed in the absence of extinction. Typically, one measures the extinction in a line-of-sight toward a star and I_λ° is the specific intensity of the starlight. For this situation, m is the observed magnitude and $m_\circ$ the intrinsic magnitude of the stellar light at wavelength λ. If the optical depth in the ISM in between us and the star is τ_λ, the observed intensity will be given by (see Eq. 2.42)

$$I_\lambda = I_\lambda^\circ \, e^{-\tau_\lambda}, \tag{12.12}$$

hence

$$A_\lambda = -2.5 \log \left[I_\lambda / I_\lambda^\circ \right] = -2.5 \log \left[e^{-\tau_\lambda} \right] = 1.086 \, \tau_\lambda. \tag{12.13}$$

The extinction measured in magnitudes is proportional to – and almost equal to – the optical depth at the chosen wavelength. In most cases we can only measure the flux, and not the specific intensity. In that case one may replace I_λ with $\mathcal{F}_\lambda$ in the above equations, and again obtain the result Eq. (12.13).

The optical depth in dust along a path length ds may be written as (see Eqs. 2.34 and 2.39)

$$d\tau_\lambda = \sigma_\lambda \, n_\mathrm{d} \, ds = \pi a^2 \, Q_\mathrm{ext}(\lambda) \, n_\mathrm{d} \, ds, \tag{12.14}$$

where n_d is the density of dust grains, a is the grain radius, σ_λ is the cross section per dust grain and Q_ext is the *extinction efficiency*, i.e. the cross section normalized to the projected grain surface. We will return to Q_ext in Sect. 12.3.1. Substitution in Eq. (12.13) yields

$$A_\lambda = 1.086 \, \pi a^2 \, Q_\mathrm{ext} \int_0^d n_\mathrm{d}(s) \, ds = 1.086 \, \pi a^2 \, Q_\mathrm{ext} \, N_\mathrm{d}, \tag{12.15}$$

where N_d is the column density of the dust particles in the line of sight towards a background source at distance d.

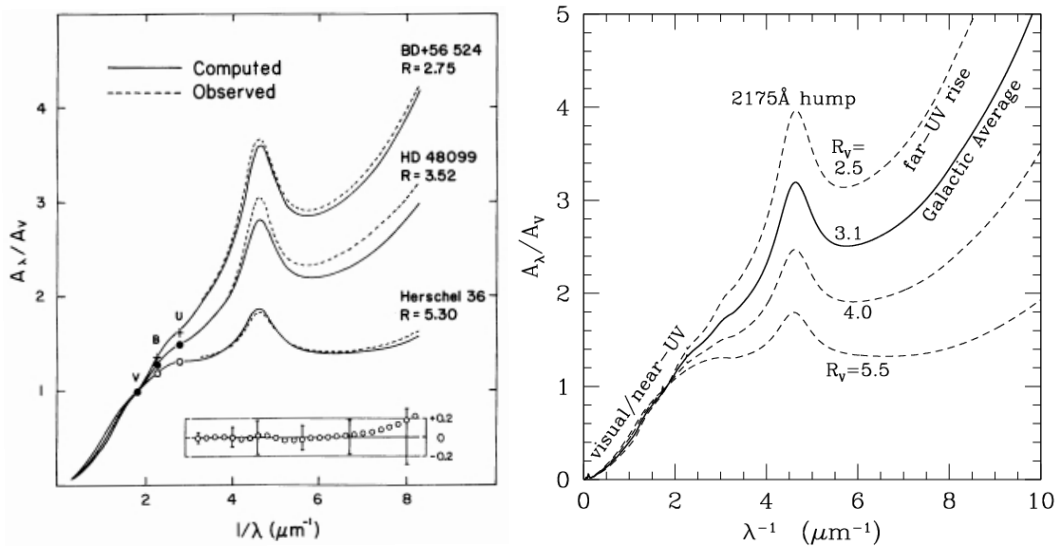


Figure 12.2: *Left: The measured wavelength dependence of the interstellar extinction for three directions into the ISM (dashed lines). The extinction is normalized to the value in the V band (5500 Å). For each direction the best fit value R_V and the name of the star that is targeted is given. The extinction in the direction of Herschel 36 – the ionizing star in the H II region M8 – is considered "peculiar". The full lines are fits to the measured A_λ/A_V. Right: The Galactic average of the interstellar extinction has $R_V = 3.05 \pm 0.15$; often 3.1 is adopted.*

Interstellar extinction law

The wavelength dependence of the interstellar extinction in the direction of three stars is shown in Fig. 12.2, where $A_\lambda/A_V = Q_{\rm ext}(\lambda)/Q_{\rm ext}(V)$ is the interstellar extinction normalized to the photometric V band. The functional behaviour of this quantity is refered to as the *interstellar extinction law*. It only depends on the (mean) intrinsic properties of dust particles in the beam towards the star, and not on the length of the beam. The use of A_V, the extinction in the V band (see § 2.6) to normalize the extinction law is arbitrary and one could argue that it is more meaningful to use, for instance, the extinction in the I or J filter (centered around ~ 0.9 and 1.25 μm respectively) as for these wavelengths the extinction is almost independent of the direction in which we look.

A very useful parameterization of the extinction curve within the Milky Way was provided by Cardelli et al. (1989, ApJ 345, 245). They describe the entire wavelength range from ultraviolet to infrared using a fit function with seven adjustable parameters. However, in the wavelength range $3030\mathring{A} < \lambda < 3.5\,\mu$m their fit function depend only on λ and the single parameter, the *total-to-selective extinction*

$$R_V \equiv \frac{A_V}{A_B - A_V} = \frac{A_V}{E(B-V)}, \tag{12.16}$$

where A_B and A_V are the extinctions measured in the B and V photometric bands, and

$$E(B{-}V) = A_B - A_V = (B - B_\circ) - (V - V_\circ) = (B - V) - (B{-}V)_\circ \qquad (12.17)$$

is the visual selective extinction or *color excess*. Because the extinction increases from red to blue, the light reaching us from stars will be 'reddened' owing to a greater attenuation of the blue light. The quantity $E(B{-}V)$ is therefore also referred to as the *reddening*.

We may also express the properties of interstellar extinction using the reddening in a magnitude band centered at wavelength λ, $E(\lambda - V)$, normalized to the visual selective extinction

$$\frac{E(\lambda - V)}{E(B{-}V)} = \frac{A_\lambda - A_V}{E(B{-}V)} = R_V \left(\frac{A_\lambda}{A_V} - 1 \right). \qquad (12.18)$$

The curve of $E(\lambda - V)/E(B{-}V)$ versus $1/\lambda$ is also known as the *reddening law*. It shows the same behavior as those of A_λ/A_V displayed in Fig. 12.2. It is however easier to measure the reddening of a star than the actual extinction A_λ. This is because the intrinsic brightness of the star is not a priori known, but its intrinsic (unreddended) color $(B{-}V)_\circ$ can be determined through the spectral type. R_V can be measured by extrapolation of the $E(\lambda - V)/E(B{-}V)$ extinction curve to $\lambda^{-1} \to 0$, i.e. $A_\lambda \to 0$. The ratio of absolute extinction A_λ/A_V can be derived from the reddening when R_V is known.

Sightlines through diffuse gas in the Milky Way have $R_V \simeq 3.05 \pm 0.15$ as an average value. The smallest well-determined value is $R_V = 2.1$ toward the star HD 210121 (Welty & Fowler 1992). Sightlines through dense regions tend to have larger values of R_V; the sightline toward HD 36982 has $R_V \simeq 5.7$ (Cardelli et al. 1989; Fitzpatrick 1999). Intuitively, we may expect that if the grains were large compared to the wavelength, the extinction cross section would equal the geometric cross section of the particle – i.e. it would be independent of wavelength with $R_V = \infty$ (as $A_B = A_V$ in Eq. 12.16). The rise of the extinction with decreasing wavelength λ implies that smaller grains than the wavelength must be making an appreciable contribution to the extinction, down to $\lambda \sim 0.1\mu$m. In this context, 'small' means (approximately) that $2\pi a/\lambda \lesssim 1$. Thus interstellar dust must include a large population of grains with radii $a \lesssim 0.015\,\mu$m = 150 Å. Figure 12.3 again shows the interstellar extinction curve, with a discussion of its shape in the caption from the perspective of grain size (and composition).

In our Milky Way the interstellar dust is concentrated in the galactic plane, with an effective scaleheight of about 100 pc. The mean reddening in the plane is about 0.61 magn kpc^{-1}. If we take $R_V = 3.1$ we get an $A_V = 1.9$ magn kpc^{-1}. Beware that in reality the distribution of dust is very patchy, i.e. concentrated in small and large interstellar clouds, and that there are directions in which the reddening deviates a factor of 5 to 10 from the mean. Measurements of the correlation between the column density of gas and the interstellar reddening yield an average gas-to-color-excess ratio

$$\frac{N(\mathrm{H\,I} + \mathrm{H_2})}{E(B{-}V)} \quad = \quad 5.8 \times 10^{21} \quad \mathrm{atoms\,cm^{-2}\,mag^{-1}}, \qquad (12.19)$$

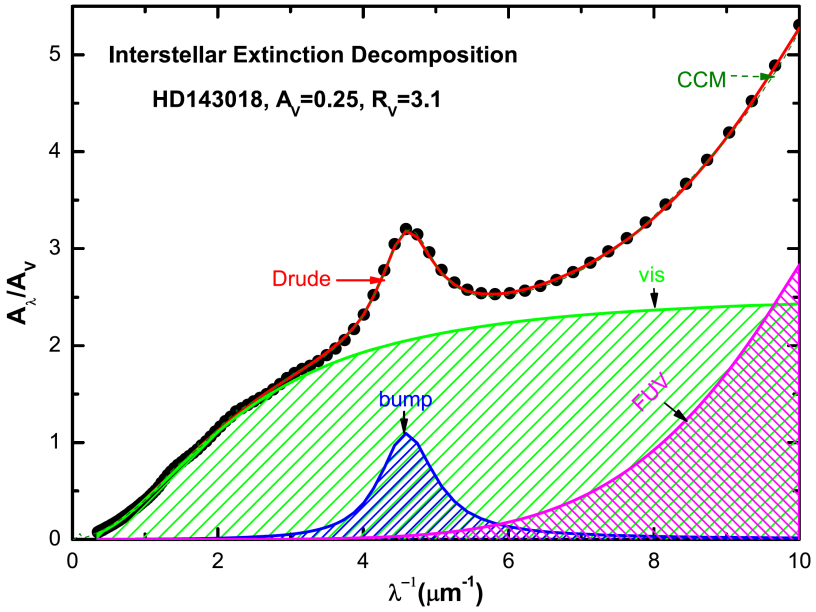


Figure 12.3: *Decomposition of the interstellar extinction curve (black circles) observed toward HD 143018 into three components: a visual component (green line), a bump component (blue line), and a far-UV component (purple line). The solid red line labelled 'Drude' is the sum of these three components. The dashed green line is the Cardelli, Clayton, & Mathis (1989) fit for* $R_V = 3.1$*. The physical basis for this decomposition lies in the interstellar grain size distribution (Greenberg 1978). The near-IR/visual part is dominated by silicate and graphite grains with* $a > 250$ *Å and saturates at* $\lambda^{-1} > 4\,\mu\mathrm{m}^{-1}$*. PAHs and ultra-small graphitic grains of* $a < 250$ *Å produce the 2175 Å bump, while ultra-small silicate grains of* $a < 250$ *Å mainly absorb at far-UV wavelengths (Xiang et al. 2017). In a comparable model by Weingartner & Draine (2001) the far-UV rise is due to both ultra-small graphitic grains and silicate grains. From: Xiang et al. 2017.*

where $N(\mathrm{H\,I} + \mathrm{H}_2) = N(\mathrm{H\,I}) + 2N(\mathrm{H}_2)$ (Bohlin et al. 1978; Rachford et al. 2009). For sight lines with $R_V \sim 3.1$, this implies that $A_V/(N(\mathrm{H}) + 2N(\mathrm{H}_2)) = 5.3 \times 10^{-22}$ mag cm^{-2} H^{-1} (see Fig. 11.2).

12.2.1 Dust density diagnostics: optical extinction

The density of dust particles ρ_d in gr cm^{-3} may be derived from the extinction at optical wavelengths. To this end we use the observed extinction in the V-band in the diffuse ISM per kpc. Building on Eq. (12.15) we may write

$$A_\lambda = 1.086\,\pi a^2\,Q_\mathrm{ext}\,n_\mathrm{d}\,D = 1.086\,\frac{3}{4}\left(\frac{Q_\mathrm{ext}}{a}\right)\frac{\rho_\mathrm{d}}{\rho_\mathrm{s}}\,D, \tag{12.20}$$

where ρ_s is the volumetric mass density of the individual grains; for silicates this is about 2.5 gr cm^{-3}. The distance D is 1 kpc for $A_V = 1.9$ magn kpc^{-1} (see above). From Fig. 12.4

we read off that at the center of the V band (at 0.55 μm), $Q/a = 10 \times 10^4$ for a in cm. Substitution yields $\rho_{\rm d} = 1.9 \times 10^{-26}$ gr cm^{-3}.

The density of hydrogen atoms in the diffuse ISM is about 1 cm^{-3} (see Sect. 5.1), therefore $\rho_{\rm H} = 1.7 \times 10^{-24}$ gr cm^{-3}. Taking into account that the hydrogen mass fraction is 0.71, we derive a dust-to-gas mass ratio

$$f_{\rm dg} = 0.71 \frac{\rho_{\rm d}}{\rho_{\rm H}} = 0.008, \tag{12.21}$$

a result consistent with the order of magnitude discussion in Sect. 5.1.

For the solar chemical abundance pattern about 1.2 percent of the species consist of elements more massive than helium. So, roughly 2/3th of these heavier elements are locked up in grains. Indeed, measurements of interstellar gas abundances of elements such as Mg and Fe show strong depletion because these elements are locked up in dust (see again Fig. 1.5).

12.3 The temperatures of interstellar grains

Hendrik van de Hulst noted — in 1946, long before observations could confirm his idea — that interstellar grains should mostly absorb at short, optical and UV wavelengths, and emit at long, far-infrared wavelengths and that their temperature would be very low. He predicted equilibrium temperatures of 10−20 K for isolated particles of diameter 0.1 μm. Such emission was indeed detected for the first time in 1973 by Judith Pipher. Dust emission turns out to be an important component ($\sim$ 10−30 %) of the total radiated luminosity of our Milky Way. Since then, several infrared satellites have made extensive studies of the infrared sky, e.g. IRAS, ISO, MSX, Akari, and recently the Spitzer Space Telescope and HERSCHEL Space Observatory. These telescopes have mapped the sky in the 10-500 μm wavelength range and have made detailed spectroscopic studies of interstellar matter.

Interstellar grains exchange energy with their surroundings through the absorption and emission of radiation. In regions of high density collisions with gas particles and exothermal chemical surface reactions may also heat the grains. In most environments radiative processes dominate. Let us first consider the simplest case: spherical grains that absorb and emit as black bodies and that are in thermal equilibrium with the interstellar radiation field.

A grain particle (or an elementary volume of gas) that is in thermal equilibrium and that absorbs and emits energy by radiative processes fulfills the constraint of *radiative equilibrium*

$$4\pi \int_0^\infty \chi_\nu(r) \left[S_\nu(r) - J_\nu(r)\right] \, d\nu = 4\pi \int_0^\infty \chi_\nu(r) \left[S_\nu(r) - \frac{c}{4\pi} u_\nu(r)\right] \, d\nu = 0. \tag{12.22}$$

The equation describes that the total amount of energy that is absorbed by an individual dust grain per second ($4\pi \int_0^\infty \chi_\nu J_\nu \, d\nu$) must be equal to the total amount of energy that is emitted from the volume in the same time interval ($4\pi \int_0^\infty \chi_\nu S_\nu \, d\nu = 4\pi \int_0^\infty \eta_\nu \, d\nu$).

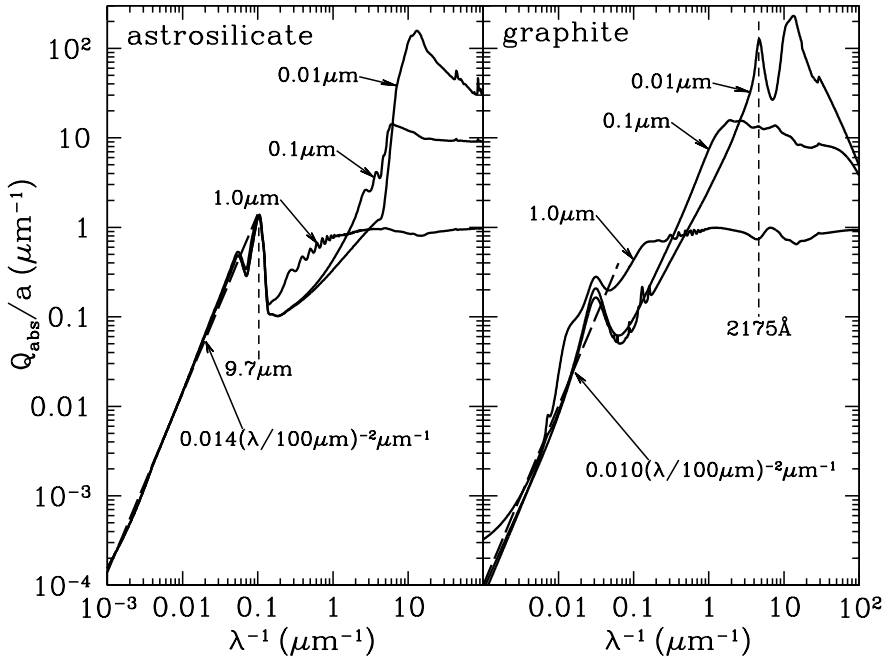


Figure 12.4: *Absorption efficiency $Q_{\rm abs}$ divided by grain radius a for spheres of amorphous silicates (left) and graphite (right). Also shown are power-laws that provide a reasonable approximation to the opacity for $\lambda \gtrsim 20\,\mu$m. From: Draine (2011).*

We express the linear extinction coefficient χ_ν as we have done above (see Eqs. 2.34 and 12.14)

$$\chi_\nu = \sigma_{\rm d}(\nu)\, n_{\rm d} = \pi a^2\, Q_{\rm ext}(\nu)\, n_{\rm d}, \tag{12.23}$$

where $n_{\rm d}$ is the density of dust grains, a is the grain radius, and $Q_{\rm ext}$ is the *extinction efficiency*. If we assume that the grains absorb and emit as black bodies, $Q_{\rm ext}(\nu)$ is unity for all frequencies (but see below). For the same reason $S_\nu = B_\nu(T_{\rm d})$, where $T_{\rm d}$ is the temperature at which the grains radiate. We find (using Eqs. 2.12 and 2.30) that

$$T_{\rm d} = \left(\frac{c}{4\sigma}\, u\right)^{1/4} \simeq 3.2\ {\rm K}, \tag{12.24}$$

where we have adopted for the total energy density of starlight $u = 0.54\ {\rm eV\,cm^{-3}} = 8.65 \times 10^{-13}$ erg cm^{-3} (see Table 1.3), a result first derived by Arthur Eddington in 1926. The temperatures of the dust grains in the diffuse ISM are actually more like 15-20 K. To explain the reason for this, we need to take a closer look at the extinction efficiency.

12.3.1 Extinction efficiency

The extinction efficiency is in reality not gray (i.e. independent of wavelength), as we have assumed above. Moreover, it consists of two contributions, one due to absorption and one due to scattering.

$$Q_{\rm ext}(\nu) = Q_{\rm abs}(\nu) + Q_{\rm sca}(\nu). \tag{12.25}$$

The scattering efficiency does not play a role in the thermal balance of the dust – scatterings only change the direction of the incident photons, hence do not contribute in any way to the heating or cooling of the particles.

It is beyond the scope of the lectures to discuss the theory that describes the way particles interact with radiation. Pioneering work in this field has been performed by Gustav Mie (1908) and Peter Debye (1909) for spherical particles. Codes are freely available on the web that calculate the extinction efficiencies for particles based on this *Mie theory*. Noteworthy to mention though is that the extinction efficiencies are a function of two parameters: (1) a dimensionless size parameter

$$x = \frac{2\pi a}{\lambda}, \tag{12.26}$$

and (2) the complex refractive index of the grain material

$$m = n - ik, \tag{12.27}$$

describing the composition of the particles. k determines the absorption cross-section and n is a measure of the scattering. In order to calculate the extinction curves for an assumed grain composition the real and imaginary part of the complex refractive index must be known. These quantities k and n, sometimes misleadingly referred to as 'optical constants', in general depend on wavelength.

In Mie theory it is found that for values x up to $\sim$3, the extinction increases proportional to x and for larger values of x the curve oscillates. For grains with a range in sizes the extinction will increase until $x \sim 3$ or $1/\lambda \sim 1/(2a)$ and for larger grains will remain roughly constant. The extinction is largest at a wavelength roughly equal to the size of the particle. For grains much larger than the wavelength the extinction is almost constant (and of the order of the geometric cross-section). These basic trends can be seen in the more detailed calculations of the silicate and graphite grains shown in Fig. 12.4. The observed extinction curve of interstellar dust actually never reaches a constant value and so this must mean that very small particles exist in interstellar space and that there are more small than large grains.

Let us now return to the question as to why the temperature of dust is $\sim$10$-$15 K, rather than 3 K. Figure 12.4 shows $Q_{\rm abs}/a$ – which is proportional to the absorption cross section per unit volume – for graphite and silicate spheres with radii $a = 0.01, 0.1$, and 1μm. We focus on grains of 0.1 μm, for which $Q_{\rm abs} \sim 1$ at ultraviolet wavelengths but for which $Q_{\rm abs}$ becomes very small at infrared wavelengths. When the wavelength is much larger than the dimension of the dust grain, radiation hardly 'notices' its presence: no radiation is absorbed. Because

of Kirchhoff's law (Eq. 2.58), which states that the emission is proportional to the extinction efficiency, also hardly any radiation is emitted. For the grain to be in thermal equilibrium, i.e.

$$\int_o^\infty Q_{\rm ext}(\nu)\, u_\nu\, d\nu = \frac{4\pi}{c} \int_0^\infty Q_{\rm ext}(\nu)\, B_\nu(T_{\rm d})\, d\nu, \tag{12.28}$$

the temperature $T_{\rm d}$ must be higher (than $\sim$ 3 K): to compensate for the efficient heating of the grain by optical and ultraviolet light, the poor cooling efficiency of the grains at infrared wavelengths requires a higher grain temperature.

When the particles are much smaller than the wavelengths some approximate formulas can be derived for $Q_{\rm sca}$ and $Q_{\rm ext}$. The wavelength dependence of $Q_{\rm abs}(\lambda)$ is often approximated by a power law, so $Q_{\rm abs}({\rm FIR}) \propto \lambda^{-\beta}$. β depends on the nature of the dust material and has a value between 1 and 2. For weakly absorbing materials with $\beta = 1$ van de Hulst showed in 1946 that $T_{\rm d} \sim 15$ K. Strongly absorbing dust species such as graphite reach equilibrium temperatures that can be a factor two higher than those of silicates. These numbers hold for the diffuse ISM. For $Q_{\rm sca}(\lambda)$ the wavelength dependence will be λ^{-4}.

Depending on local conditions, dust temperatures can vary significantly from the ones derived above for the diffuse ISM. For instance, towards the galactic centre the ISRF intensifies and so the dust temperature will increase. In the cores of molecular clouds the diffuse ISRF is highly attenuated and the temperature of silicates can drop to 7 K when shielded by A_V = 10 mag in the centre of the cloud. Close to early-type stars (that are hot) dust can be much hotter due to the much stronger radiation field. Circumstellar dust shells generally show the hottest dust, with values up to the sublimation temperature (1000-1700 K, depending on dust species and gas pressure).

12.4 Temperatures of very small grains

In radiative equilibrium (see above) the grain is at this steady-state or equilibrium temperature and the vibrational energy content of the particle is

$$E_{\rm vib}(T_{\rm eq}) = \int_0^{T_{\rm eq}} C(T)\, dT, \tag{12.29}$$

where $C(T)$ (in erg K^{-1}) is the *heat capacity* of the grain at temperature T. If the particle is so small that the energy content becomes small relative to the mean energy per absorbed photon, i.e. $E_{\rm vib} \lesssim \langle h\nu \rangle_{\rm abs}$, then individual photon absorptions will cause pronounced upward jumps in the grain temperature. In between photon absorptions substantial radiative cooling of the grain will take place. As a result, the grain temperature T will be a strongly fluctuating quantity, with large excursions above and below $T_{\rm eq}$.

Figure 12.5 shows the temperature histories of five graphitic grains over the span of about a day. For the grain sizes shown here, $Q_{\rm abs}(\lambda) \propto a$, so that the absorption cross section $\sigma_{\rm d}(\lambda) =$

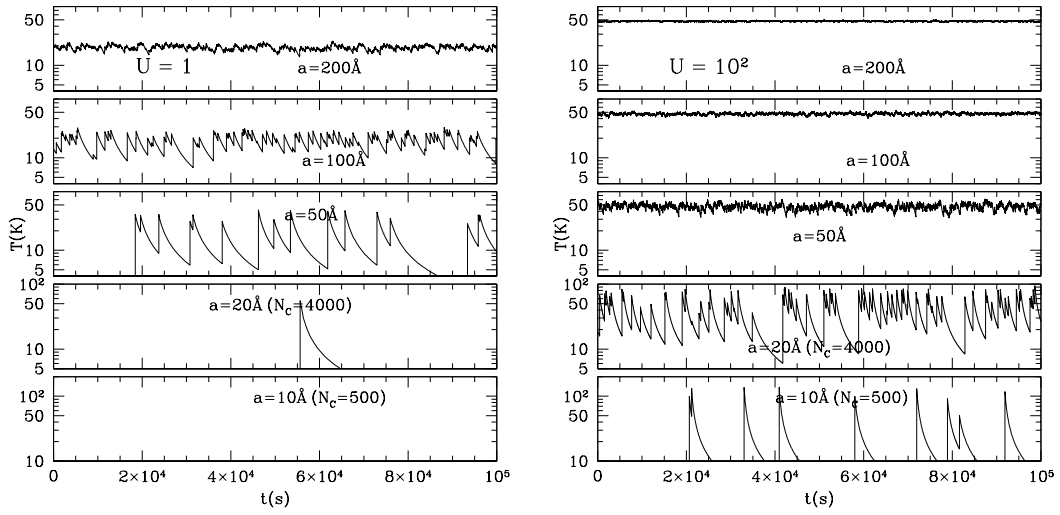


Figure 12.5: *Temperature evolution as a function of time over a period of* 10^5 *s* (~ 1 *day) for carbonaceous grains of different radii in two radiation fields: the local starlight intensity* $U = 1$ *(left panel) and a hundred fold the local starlight intensity* $U = 10^2$ *(right panel). For the definition of* U *see Eq. (6.8). The importance of quantized stochastic heating is most pronounced for the smallest sizes. For these grains the number of network formers* $N_{\rm C}$ *(in this case single carbon atoms) is given as well. Notice that the typical cooling time is fairly independent of particle size and about* 10^4 *s – best seen in the* $a = 20\,\AA$ *and* $U = 1$ *panel. From: Draine (2011).*

$\pi a^2\,Q_{\rm abs}(\lambda) \propto a^3$. From Fig. 12.4 we may estimate that for very small carbonaceous grains, $Q_{\rm abs} \simeq 2\,a$ [μm] = $2 \times 10^4\,a$ [cm] at a mean frequency of the photons responsible for the heating of $\overline{\nu} = 3 \times 10^{14}$ Hz (or 1 μm). The rate (per second) at which these particles absorb incident photons is given by

$$\begin{aligned} R_{\rm abs} &= 4\pi \int_0^\infty \sigma_{\rm d}\,\frac{J_\nu}{h\nu}\,d\nu = 4\pi \int_0^\infty \pi a^2 Q_{\rm abs}\,\frac{c}{4\pi} U \frac{u_\nu^{\rm MMP83}}{h\nu}\,d\nu \\ &\simeq 2 \times 10^4\,\pi a^3\,\frac{c}{h\overline{\nu}}\,U\,u^{\rm MMP83} \approx 1 \times 10^{-6}\,U\,\left(\frac{a}{10\,\mathring{A}}\right)^3, \end{aligned} \tag{12.30}$$

where we have adopted for the total energy density of the mean interstellar radiation field the value 1.05×10^{-12} erg cm^{-3} from Table 6.1. The constant U is defined in Eq. (6.8). The typical time between two successive absorptions of a photon is thus

$$\tau_{\rm abs} \approx 1 \times 10^6\;U^{-1}\,\left(\frac{a}{10\,\mathring{A}}\right)^{-3}\;{\rm s}. \tag{12.31}$$

For a 50 Å grain and $U = 1$ we find $\tau_{\rm abs} \approx 10^4$ s.

The time to cool below ∼5 K is independent of grain size for $a \lesssim 200$ Å and about

$$\tau_{\rm rad} \approx 10^4 \ {\rm s} \tag{12.32}$$

The estimate shows that for grains smaller than about 50 Å, $\tau_{\rm rad} < \tau_{\rm abs}$. When a photon absorption does take place, the small heat capacity of the grain results in a high peak temperature. It is clear that one cannot speak of a representative grain temperature under these conditions – one must instead use a temperature distribution function. For grains larger than about 50 Å, $\tau_{\rm rad} > \tau_{\rm abs}$. Though absorption events occur more frequently, the temperature rise at each event is reduced by the increased heat capacity, and the temperature varies over only a small range, as seen for the $a = 200$ Å grain in Fig. 12.5.

Radiation from 'super-heated' dust grains has first been detected using IRAS. Apart from the expected thermal emission from 'large' interstellar dust grains at temperatures of 15-20 K with a peak in the flux near 100 μm wavelength, the IRAS satellite also observed strong emission at shorter wavelengths that cannot be explained with these large cold grains. The explanation for this extra emission at short wavelengths is of course very small grains, with sizes as small as large molecules. This small particle population is dominated by PAHs, as later shown through spectroscopy with ISO and Spitzer. PAHs and nano-grains are transiently heated to a temperature of typically a few 100 K and emit photons at that temperature, before cooling down vary rapidly (within seconds) to much lower temperatures. The maximum temperature that small PAH molecules may reach by means of this quantum heating process is about 2000 K, when exposed to an intense radiation field that emits FUV photon with an energy approaching 13.6 eV.

12.5 Dust mass diagnostics: FIR continuum emission

In a typical spiral galaxy, perhaps a third of the energy radiated by stars is absorbed by dust grains and re-emitted in the infrared. The spectrum of this emission is determined by the temperatures and composition of the dust grains.

The observed far-infrared emission of interstellar dust can be used to derive the total mass of interstellar grains, provided we make some assumptions about the grain properties. Once we have the total grain mass, we can derive the total gas plus dust mass by using the dust to gas mass ratio (see e.g. Eq. 5.3, though more often a ratio of 1/150 is adopted). This kind of calculation is often done in the literature, but it has some important uncertainties, related to the assumptions usually made. An important one is that of an isothermal dust cloud. This condition may be met in the diffuse ISM, but will certainly break down in dense molecular clouds. Sub-millimeter radiation probes the innermost and coldest regions of such clouds, where the bulk of the dust mass is and that can be assumed roughly isothermal. Hence, the method can still be applied in these cases.

Let us assume an optically thin dust cloud. The flux that we receive from this cloud is (using Eq. 2.34)

$$\mathcal{F}_\lambda = \frac{4\pi\,\eta_\lambda\,V}{4\pi\,d^2} = \frac{\chi_\lambda B_\lambda(T_\mathrm{d})\,V}{d^2} = \frac{\chi'_\lambda \rho_\mathrm{d} B_\lambda(T_\mathrm{d})\,V}{d^2}, \tag{12.33}$$

where ρ_d is the dust density and V the volume of the dust cloud. In the second equality we have used the Kirchhoff-Planck relation Eq. (2.58). d is the distance to the cloud. We find for the dust mass

$$M_\mathrm{d} = \rho_\mathrm{d} V = \frac{\mathcal{F}_\lambda\,d^2}{\chi'_\lambda B_\lambda(T_\mathrm{d})}. \tag{12.34}$$

Alternatively, using Eq. 12.23

$$\mathcal{F}_\lambda = \frac{4\pi\,\eta_\lambda\,V}{4\pi\,d^2} = \frac{\chi_\lambda B_\lambda(T_\mathrm{d})\,V}{d^2} = \frac{\pi a^2}{d^2}\,Q_\lambda\,B_\lambda(T_\mathrm{d})\,N, \tag{12.35}$$

where N is the total number of spherical dust grain of uniform size, composition and temperature in the cloud. a is the radius of these dust grains. The dust mass is given by

$$M_\mathrm{d} = \frac{4\pi}{3}\,a^3\,\rho_\mathrm{s}\,N \tag{12.36}$$

where ρ_s is the mass density of the dust particles. We may eliminate N from Eqs (12.35) and (12.36) and rearrange terms such as to obtain an expression for the total dust mass. This yields

$$M_\mathrm{d} = \frac{4\rho_\mathrm{s}\,\mathcal{F}_\lambda\,d^2}{3B_\lambda(T_\mathrm{d})}\left(\frac{a}{Q_\lambda}\right) \tag{12.37}$$

Figure 12.4 shows that at long wavelengths the plotted ratios a/Q_λ do not depend on grain size – a consequence of the grains being small compared to the wavelength at which they radiate. If the temperature is known (e.g. using the cold gas temperature diagnostics described in Sect. 4.4) a flux measurement at a single wavelength suffices to compute the total dust mass. If T_d is not known, we may measure the flux at several wavelengths and constrain the temperature fitting the overall behavior of the energy spectrum.

12.5.1 Dust to gas ratio diagnostics: FIR continuum emission

We can yet again derive the gas to dust mass ratio, as was done in Sect. 12.2.1. This time we do not use the *extinction* at short wavelengths, but rather the thermal *emission* at infrared wavelengths. The infrared emission method is however more difficult to use than is the extinction method.

For optically thin dust in the line of sight, the flux observed in a radius of angle θ at wavelength λ is

$$\mathcal{F}_\lambda = I_\lambda\,d\omega\ = \eta_\lambda\,D\,\pi\theta^2 = S_\lambda\,\tau_\lambda\,\pi\theta^2 = B_\lambda(T_\mathrm{d})\,\tau_\lambda\,\pi\theta^2, \tag{12.38}$$

where we have used Eq. (2.14) to express the flux in specific intensity. D is the line of sight path length. For an adopted dust temperature, the only unknown in this equation is τ_λ and

so the optical depth of the dust grains can be determined. However, when the adopted $T_{\rm d}$ is too low, the optical depth will be larger than computed. Clearly, the accuracy of the method critically depends on our knowledge of $T_{\rm d}$.

Borrowing from Eq. (12.20), we may write

$$\mathcal{F}_\lambda = B_\lambda(T_{\rm d})\,\pi\theta^2\,\frac{3}{4}\left(\frac{Q_{\rm ext}}{a}\right)\frac{\rho_{\rm d}}{\rho_{\rm s}}\,D. \tag{12.39}$$

This results in a value for the column density of the dust, $\rho_{\rm d}\,D$. If for the same line of sight we measure the column density in hydrogen, $\rho_{\rm H}\,D$ (or the total hydrogen column density $N_{\rm H} = \rho_{\rm H}\,D/m_{\rm H}$), we may apply Eq. (12.21) to obtain the dust-to-gas ratio $f_{\rm dg}$.

Exercise 12.1

In Section 12.5 it is stated that if the temperature of the dust in a (dust) cloud is not known, one may measure the flux at several wavelengths and constrain the temperature fitting the overall behavior of the energy spectrum.

We consider dust grains for which the absorption efficiency $Q_{\rm abs}$ is given by (see Fig. 12.4)

$$\frac{Q_{\rm abs}}{a(\mu{\rm m})} = 0.014 \left(\frac{\lambda}{100\,\mu{\rm m}} \right)^{-2} \qquad \mu{\rm m}^{-1}. \tag{12.40}$$

This relation is valid for FIR wavelengths. As interstellar dust+gas clouds are very cold, it is reasonable to assume that $hc/\lambda kT \ll 1$ for this wavelength regime. We also assume the dust grains emit radiation following a Planck function.

a) Show that for an optically thin cloud the flux behaves as

$$\mathcal{F}_\nu(\lambda) \propto \lambda^{-6}$$

The implication is that far in the infrared such clouds may become undetectable due to sensitivity limits of the instrument.

b) Would it be useful to measure several flux points in this regime in order to constrain the dust temperature?

Exercise 12.2

a) Show that for a head-on elastic collision of a projectile particle on a target particle the transferred energy is given by Eq. (12.5).

b) Why, for given m_p and m_t, is this the maximum energy that can be transferred in an elastic collision?

c) For what mass ratio m_p/m_t is the transferred energy maximal?

Exercise 12.3

This could be a nice exam question. Sputtering acts to erode grains at a rate (see e.g. Eq. 12.6)

$$\frac{{\rm d}a}{{\rm d}t} = -\beta \cdot n_{\rm H}, \tag{12.41}$$

independent of a. We assume that β, often expressed in $\mathrm{cm}^3\,\mu\mathrm{m\,yr}^{-1}$, is constant. n_H is the density of hydrogen particles. Suppose that the grain-size distribution at $t = 0$ is a power-law

$$\frac{1}{n_\mathrm{H}}\frac{\mathrm{d}n}{\mathrm{d}a} = \frac{A_0}{a_\mathrm{max}}\left(\frac{a}{a_\mathrm{max}}\right)^{-p} \qquad 0 \le a \le a_\mathrm{max}. \tag{12.42}$$

The maximum grain size is a_max.

a) After how long have all grains been eroded? Express this time t_max in terms of the given quantities.

b) Let V_0 be the initial volume of grain material per hydrogen nucleus. Express V_0 in terms of A_0, a_max, and p.

c) Obtain an algebraic expression for $V(t)/V_0$ in terms of $y \equiv \Delta a/a_\mathrm{max} = \beta\, n_\mathrm{H}\, t/a_\mathrm{max}$ and p. Δa is the decrease in grain radius of particles that still exist after time t.

Exercise 12.4

We build on the results from the previous problem and consider sputtering in the hot coronal gas (see e.g. Table 1.1). We assume $p = 3.5$, and $a_\mathrm{max} = 0.3\,\mu$m. Assuming a temperature of 10^6 K, we find from Eq. (12.6) that $\beta = 0.5 \times 10^{-6}\ \mathrm{cm}^3\,\mu\mathrm{m\,yr}^{-1}$. For the hydrogen particle density we adopt $n_\mathrm{H} = 0.01\,\mathrm{cm}^{-3}$

a) Make a plot of $V(t)/V_0$ s a function of $\Delta a/a_\mathrm{max}$.

b) Graphically (or, alternative, using a root-finding routine) estimate $\Delta a/a_\mathrm{max}$ such that $V/V_0 = 1/2$.

c) What time t is required to sputter away 50% of the mass in grains?

d) What time t is required to sputter away all of the mass in grains?