Radiative processes and their application to the intergalactic and interstellar medium

Tom Theuns[1] & Michele Fumagalli[2]

[1]tom.theuns@durham.ac.uk, Institute for Computational Cosmology, Durham, UK

[2]michele.fumagalli@unimib.it, University of Milano Bicocca, Italy

Introduction

Most of the physical properties of galaxies is inferred from their spectra (from composition of their stars and their gas, to their star formation rates). Therefore understanding how the emission of light by stars (and AGN), combined with the emission and absorption of light by gas and dust, determines the light detected from a galaxy is a skill worth acquiring. These lecture notes treat some of the main physical processes through which matter and radiation interact, with applications to the physics of the intergalactic medium (IGM, the stuff between the galaxies) and the interstellar medium (ISM, the stuff between stars in a galaxy).

A galaxy such as the Milky Way would run out of gas to form stars in about 1 Gyr. Fresh fuel for star formation is thought to originate from stellar evolution (stellar mass-loss) and from cosmological accretion from the IGM and cooling of gas in the dark halo hosting the galaxy. The rate at which a galaxy accretes mass increases strongly with redshift, which is the basic reason why high-z galaxies form stars so vigorously. But how much gas is there in the IGM, what are its properties, and how can we study these? Because of its relatively low density, the IGM is most easily studied in absorption by observing a bright background source - for example a quasar. Analysing the intervening absorbers has lead to many surprising discoveries. For example the IGM is highly ionised - presumably due to ionising radiation emitted by galaxies and quasars. But the IGM also contains metals - elements synthesised in stars - which is rather unexpected and implies that galaxies also lose mass (at least metals) to the IGM. With the physics of the IGM relatively simple, it is also a laboratory for testing physics - from measuring the baryonic acoustic oscillation scale up to $z > 2$, to putting limits on the mass of the neutrino and on the nature of the dark matter, and even setting limits on the constancy of the fine-structure constant (as well as of other physical 'constants').

The interstellar medium (ISM) is the *stuff* between the stars, and consists of *gas* (from highly ionised over neutral to molecular), *dust* (solid particles), *radiation*, *magnetic fields*, and *cosmic rays*. These different components interact with stars and the large-scale dynamics of the spiral disk in many complex and not well-understood ways. For example star formation is made possible because the molecules and dust allow gas to cool by converting thermal energy to radiation. This is a crucial step in the huge increase in gas density from typical ISM densities of ~ 1 hydrogen atom per cm^{-3} to the stellar density of $\geq 10^{23}$ cm^{-3}. The dust grains are important in the chemistry of the ISM, regulating the formation of molecules. But the dust grains themselves, and the elements they are made of, were synthesised in stars. Just like the chicken and egg problem you may wonder what came first: dust and molecules (to make stars), or stars that produced the elements that dust and molecules are made of.

When the gas in a galaxy cools very rapidly its disk may become unstable to star formation. The ensuing starburst and the associated powerful super novae explosions, can then blow most of the ISM out of the galaxy. Star formation is then halted, until the galaxy has accreted enough material from its surroundings to become active again. Such bursty behaviour is seen in numerical simulations of small galaxies, but also observationally when two larger galaxies merge. In the Milky Way, star formation and the feedback from the associated super novae, seem to self regulate - although why this would be so is not particularly well understood - and the MW is forming stars at a very modest rate of a about 1-2 M$_\odot$ yr^{-1}, as compared to the $10^2 - 10^3$ M$_\odot$ yr^{-1} of starburst galaxies. Yet star formation may not have always been so gentle in the Milky Way: as compared to the Universe as a whole, the Milky Way has lost a significant fraction of its baryons (compared to its dark matter mass), suggesting the MW was much wilder in its youth.

Interestingly several key processes in the physics of the ISM are far from well understood. Although it is well established that most, if not all, MW stars form in the Giant Molecular Clouds (GMCs) that line its spiral arms, how the GMCs themselves form, and the physics that governs them, is hotly contested. Some models of GMCs have them as long-lived entities held together by self-gravity, other models see them as transient structures swept-up by the spiral density wave. The importance of magnetic fields, and cosmic rays, is unclear. Curiously, the energy density of the gas, magnetic field, and cosmic rays, are very similar: what is the origin of this equipartition? And does it matter? Is the ISM of other galaxies similar to that of the MW or

not?

The complexity of the ISM should not be an excuse to shy away from its study. In particular if we want to study how stars form under different conditions, for example in other galaxies, or at very high redshifts, we should probably make sure we understand these processes locally, where the observations are much more constraining. Most of these constraints come from spectral analysis, where we can study the gas and dust in absorption or emission. *Therefore the interaction of matter with radiation is the central theme of these notes.* It is by understanding the relation between density, temperature, abundance, radiation field, and the associated spectral signatures, than we can try to constrain the physical and chemical properties of the ISM. This will provide the data that any successful ISM model should reproduce.

The study of the ISM has surprisingly personal dimensions. Most of the mass of your own body (say in terms of O, C and Ca) was of course synthesised in stars, and cycled through the ISM before it became part of you. So: how many stars then did it take to make you? And when did these stars form, and die? The ISM also contains surprisingly complex molecules, from the popular C_2H_5OH to the much more complex RNA: may be the extreme physical conditions of the ISM (very low densities as compared to the lab, but high radiation fields) allow the fabrication of such building blocks of life? And finally is earth not some rather larger-than-average dust grain?

These lecture concentrate more on the basic physics of the interaction of light and matter, than on its application to IGM and ISM. The aim is to provide the interested reader with the background knowledge and tools to tackle reading and appreciating the subtleties in papers that interpret spectra in terms of physical properties of gas and dust. The course also involves experimenting with CLOUDY, a numerical code that solves for the detailed interaction between light and matter in situations far beyond what is possible with pen and paper. We - the authors - are keen to get constructive feedback and suggestions on these notes.

More reading: these notes rely heavily on Rybicki, G. B., & Lightman, A. P. 1986, *Radiative Processes in Astrophysics*, Sun Kwok, *Physics and Chemistry of the Interstellar Medium*, and J. E. Dyson & D. A. Williams, *The Physics of the Interstellar Medium*.

Chapter 1

Fundamentals of radiative transfer

- Definitions of flux, intensity and their relation to energy density and radiation pressure
- Concept of light ray
- Proof of the constancy of intensity along a ray, and its relation to $1/r^2$ fall-off from a source
- Derivation and applications of the equation of radiative transfer

Further reading: *Rybicki & Lightman, Chapter 1*

1.1 EM-radiation

Electro-magnetic radiation is carried by photons, for which wavelength[1] (λ, $[\lambda]$= cm) and frequency (ν, $[\nu]$=Hz) are related by

$$\lambda = \frac{c}{\nu}, \tag{1.1}$$

where by definition the speed of light $c = 2.99792458 \times 10^{10}$ cm s^{-1} in vacuum. Regions of the spectrum are traditionally divided into wavelength ranges as in Table 1.1. The quantum (photon) nature of the radiation can often be

[1] We will use the notation $[X]$ to denote the dimension of quantity X.

$\log(\lambda/\mathrm{cm})$	≤ -8.9	-6.1	-4.3	-4.1	-1.3	≥ -1.3
name	γ-rays	X-rays	UV	Visible	IR	Radio

Table 1.1: Typical wavelength ranges for various EM radiation

neglected[2], in which case we describe radiation using the concepts of flux and intensity, described below. The connection with Maxwell's equations is discussed later.

1.2 Radiative flux

Luminosity
The luminosity L ($[L]=$ erg s^{-1}) of a source is the total amount of EM-energy it radiates per unit time. Only *bolometers* can detect EM-radiation (almost) independent of λ, therefore more useful is the luminosity emitted in a given wavelength range $L(\nu)$ ($[L(\nu)]=$ erg s^{-1} Hz^{-1}]), for example in the B-band, L_{B}.

Flux
Consider a sphere of radius R centered on a source with luminosity L. The energy per unit time passing through the surface area $4\pi R^2$ is L. The flux F is the energy per unit area per unit time, so $F = L/4\pi R^2$, or in the B-band, $F_{\mathrm{B}} = L_{\mathrm{B}}/4\pi R^2$. In general the EM energy passing through any area dA per unit time is

$$dE = F dA dt\,, \tag{1.2}$$

where $[F]=$ erg s^{-1} cm^{-2}.

Intensity
Flux is the energy carried by all *rays* through an area. Construct an area dA normal to the ray, and consider all rays passing through dA within a small solid angle $d\Omega$. The energy dE passing through dA within this solid angle (per unit time, per unit frequency) is

[2]Not, for example, when photon counting affects the level of the noise in a CCD detector.

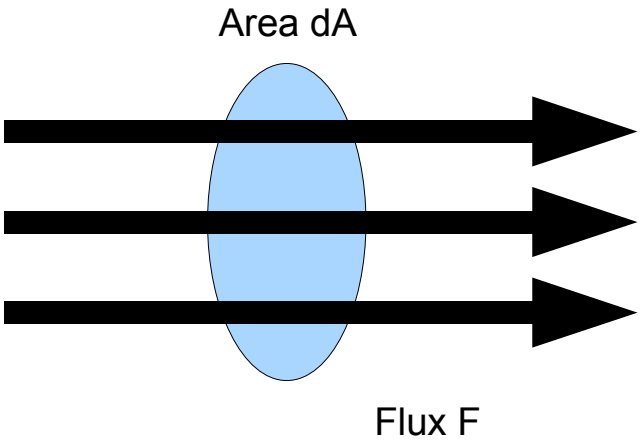

Figure 1.1: The flux F is the energy passing through the area dA per unit time, per unit area.

$$dE = I_\nu \, dA \, dt \, d\nu \; d\Omega \, , \tag{1.3}$$

which defines the *specific intensity* or *brightness* I_ν ($[I_\nu] = \mathrm{erg\ s^{-1}\ cm^{-2}\ ster^{-1}\ Hz^{-1}}$), see Fig. 1.2. If dA makes an angle θ with the direction of the ray, then $dE = I_\nu \cos\theta \; dA \, dt \, d\nu \; d\Omega$. In general I depends on position, time, direction, and frequency: we will usually suppress this dependence not to overburden notations. Re-writing this equation yields the following relation

$$I_\nu d\nu = \frac{dE/dt/dA}{d\Omega} = \frac{dF_\nu}{d\Omega} d\nu \, . \tag{1.4}$$

This shows that $I_\nu d\nu$ is the flux received per unit solid angle - in the cosmology lectures we defined this to be the surface brightness (of a source). The intensity $I_\nu d\nu$ of a light ray is therefore nothing else than surface brightness, and the fact that it is conserved along a ray, as we will demonstrate soon, is equivalent to stating that surface brightness is independent of distance.

In case I is independent of direction (*isotropic radiation*), demonstrate as an exercise that the net flux $\propto \int \cos\theta \, d\Omega = 0$. The *net* flux passing through any surface is then zero, basically because for any direction $\hat{n}$ there is as much radiation passing the area dA in the $+\hat{n}$ direction as in the $-\hat{n}$ direction.

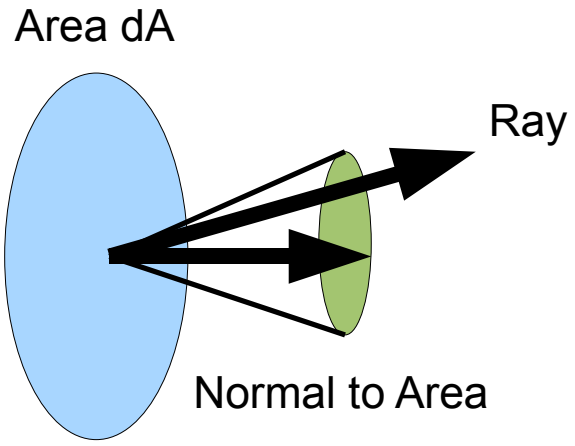

Figure 1.2: The energy carried by a ray in the solid angle $d\Omega$ perpendicular to the area dA.

Momentum flux
Since E/c is the momentum of a photon with energy E, the momentum carried through dA by a ray with intensity I, is

$$p_\nu = \frac{1}{c} \int I_\nu \cos^2\theta \, d\Omega \,, \tag{1.5}$$

where as before the ray strikes dA under the angle θ. We will later use the symbol p to denote the radiation pressure: beware.

Energy density
Consider a cylinder with radius R, base area $dA = \pi R^2$ and length $l = c\,dt$. The radiation energy is moving perpendicular to dA with speed c, therefore all energy crosses dA in time dt and

$$dE_\nu = u_\nu \, dA \, c \, dt \, d\nu \, d\Omega = I_\nu \, dA \, dt \, d\nu \, d\Omega \tag{1.6}$$

where

$$u_\nu(\hat{\Omega}) = \frac{I_\nu}{c} \tag{1.7}$$

is the *energy density* of radiation moving in direction perpendicular to dA.

Integrating over the full solid angle yields

$$u_\nu = \int u(\hat{\Omega})\, d\Omega = \frac{1}{c}\int I_\nu\, d\Omega \equiv \frac{4\pi}{c}\, J_\nu\,, \tag{1.8}$$

where

$$u_\nu = \frac{4\pi}{c}\, J_\nu \tag{1.9}$$

$$J_\nu = \frac{1}{4\pi}\int I_\nu\, d\Omega\,. \tag{1.10}$$

These relations define the *energy density* u_ν and the *mean intensity* J_ν.

Radiation pressure
Consider radiation with intensity I in an reflecting enclosure. When a photon with momentum $p = E/c$ reflects off the enclosure, it transfer twice its momentum to the mirror. Therefore the *radiation pressure* p_ν, is

$$p_\nu = \frac{2}{c}\int I_\nu\, \cos^2\theta\, d\Omega\,. \tag{1.11}$$

For isotropic radiation $I_\nu = J_\nu$ we integrate over half the total solid angle[3] to get

$$p_\nu = \frac{1}{3}\, u_\nu\,, \tag{1.12}$$

a relation between the pressure and energy density of radiation.[4]
Constancy of I along a ray in empty space
Consider two areas, $dA_{1,2}$, separated by distance D, such that all rays pass through both areas. Energy conservation guarantees that the energy dE_1 passing through dA_1 also passes through dA_2:

$$dE_1 = I_1\, dA_1\, dt\, d\Omega_1\, d\nu = dE_2 = I_2\, dA_2\, dt\, d\Omega_2\, d\nu\,. \tag{1.13}$$

However, since $d\Omega_1 = dA_2/D^2$, and $d\Omega_2 = dA_1/D^2$, we find that $I_1 = I_2$, that is *the intensity along a ray is constant* (in the absence of emission or

[3] Why only half?

[4] The ideal gas equation is $p = (\gamma - 1)\rho U$ where ρ is the mass density, and U the energy per unit mass. Therefore in this notation, $u = \rho U$ is the energy per unit volume, and $p = (\gamma - 1)u$. Comparing with Eq. 1.12 yields the familiar result that $\gamma = 4/3$ for radiation.

absorption). This does not conflict with the fact that the flux $\propto 1/D^2$: indeed, consider the flux F received from an isotropic emitter ($I = B$, some constant) with radius $R_\star$ at distance D:

$$\begin{aligned} F &= \int I \cos\theta \, d\Omega \\ &= B \int_0^{\theta_c} \cos\theta \, \sin\theta \, d\theta \int_0^{2\pi} d\phi \\ &= \pi \, B \sin^2 \theta_c \\ &= \frac{\pi B \, R_\star^2}{D^2} \, , \end{aligned} \tag{1.14}$$

where $\sin\theta_c = R_\star/D$ is the maximum angle between the surface of $R_\star$ and the direction to the centre of the emitter. This also shows that the luminosity of the emitter is $L = 4\,\pi^2 \, B \, R_\star^2$.

1.3 Radiative transfer

The fact that I is a constant along a ray, in the absence of photon interactions, whereas the flux F falls with distance to the source even in the absence of any photon interactions, makes I ideally suited to describe radiative transfer effects. The intensity of a ray can change due to *emission*, *absorption* and *scattering*. We will neglect scattering for the moment.

Emission
Consider a cylinder with radius R, base area $dA = \pi \, R^2$ and length ds. The energy produced by emission within the volume $dV = dA \, ds$ is

$$dE_\nu = j_\nu \, dV \, d\Omega \, dt \, d\nu = dI \, dA \, dt \, d\Omega \, d\nu \, , \tag{1.15}$$

which defines j_ν, the *emission coefficient* ($[j_\nu]$=erg cm^{-3} ster^{-1} s^{-1} Hz^{-1}). j_ν will depend on the properties of the material within the cylinder, and also defines the *emissivity* ϵ_ν

$$j_\nu = \frac{\epsilon_\nu \, \rho}{4\pi} \, , \tag{1.16}$$

where ρ is the mass density of the material. Emission then changes the intensity of a ray as

$$dI_\nu = j_\nu \, ds \, , \tag{1.17}$$

since $ds = dV/dA$.
Absorption
Consider radiation with intensity I_1 entering this same cylinder, and leaving it at the other end. In the presence of absorption, the intensity $I_2 < I_1$, since $E_1 = I_1\, dA\, dt\, d\Omega\, d\nu > E_2 = I_2\, dA\, dt\, d\Omega\, d\nu$. The amount of absorption $I_2 - I_1$ should be $\propto I_1\, ds$, which defines the *absorption coefficient* α_ν ($[\alpha_\nu] = \mathrm{cm}^{-1}$) as

$$dI_\nu = -\alpha_\nu\, I_\nu\, ds\,, \tag{1.18}$$

with $\alpha_\nu > 0$. The *optical depth*, τ_ν, is defined as

$$d\tau_\nu = \alpha_\nu\, ds = -\frac{dI_\nu}{I_\nu}\,. \tag{1.19}$$

An example is the absorption of light by solid spheres with radius r and (constant) number density n. The number of such spheres in the cylinder is $N = n\,V$, and the fraction of light blocked by them is $N\,(\pi\, r^2)/dA$, the ratio of the total area of all spheres over the base area of the cylinder. Therefore

$$\frac{I_2}{I_1} = 1 - \frac{N\,(\pi\, r^2)}{dA} = 1 - n\, ds\,(\pi\, r^2) = 1 + \frac{dI}{I_1}\,, \tag{1.20}$$

hence for such spheres

$$\begin{aligned} dI &= -I\, n\,(\pi\, r^2)\, ds \\ d\tau &= n\,(\pi\, r^2)\, ds\,. \end{aligned} \tag{1.21}$$

This is only valid if the spheres absorb all photons that strike them, which will be a good approximation when $r \gg \lambda = c/\nu$. When this is not true we can write $d\tau_\nu = Q(\nu)\, n\,(\pi\, r^2)\, ds$ where Q describes the ratio of the actual cross section for absorption over the geometric cross section of the sphere.

Stimulated emission, as occurs in a laser or maser, has $dI_\nu > 0$ and hence can be described as 'absorption' with $\alpha_\nu < 0$.

1.3.1 The equation of radiative transfer

When there is both emission and absorption, the equation of radiative transfer becomes

$$\frac{dI_\nu}{ds} = -\alpha_\nu\, I_\nu + j_\nu\,, \tag{1.22}$$

or in terms of the optical depth $d\tau_\nu$,

$$\begin{aligned} \frac{dI_\nu}{d\tau_\nu} &= -I_\nu + S_\nu \\ S_\nu &\equiv \frac{j_\nu}{\alpha_\nu}\,. \end{aligned} \tag{1.23}$$

S_ν, the ratio of emission coefficient over the absorption coefficient, is called the *source function*.

Emission only
For $\alpha = 0$, the solution to $dI/ds = j$ is

$$I(s) = I(0) + \int_0^s j(s')\,ds'\,. \tag{1.24}$$

Absorption only
For $j = 0$, the solution to $dI/ds = -\alpha\, I$ is

$$I(s) = I(0)\,\exp\left(-\int_0^s \alpha(s')\,ds'\right)\,. \tag{1.25}$$

General solution[5]
The formal solution for $j \neq 0$ and $\alpha \neq 0$ is

$$I(\tau) = I(\tau = 0)\,\exp(-\tau) + \int_0^\tau \exp(-(\tau - \tau'))\,S(\tau')\,d\tau'\,, \tag{1.26}$$

where s and τ are related by Eq. (1.19). Both j and α, and hence also the source function $S(s)$, will in general depend on the value of the radiation there (*i.e.* the value of I which we are trying to compute), and the equation needs to be solved numerically. In the special case where S is constant, S_0, the solution is

$$I(\tau) = I(0)\,\exp(-\tau) + S_0\,(1 - \exp(-\tau))\,. \tag{1.27}$$

[5]Demonstrate as an exercise

A system with $\tau \ll 1$ is called *optically thin*, it has $I \approx I(0)$. When $\tau \gg 1$ the system is called *optically thick*, it has $I(\tau) \approx S$. Is the Earth's atmosphere optically thin or optically thick? How about the Sun's mantle?

Mean free path
The function $\exp(-\tau_\nu)$ can be thought of as the probability that a photon travels a distance $s(\tau_\nu)$ before being absorbed. The mean optical depth at which photons get absorbed is then

$$\langle \tau_\nu \rangle = \int_0^\infty \exp(-\tau_\nu)\, \tau_\nu \, d\tau_\nu = 1\,. \tag{1.28}$$

The mean distance, $\langle s_\nu \rangle$, the photon travels follows from $\langle \tau_\nu \rangle = \alpha_\nu \langle s_\nu \rangle = 1$, hence

$$\langle s_\nu \rangle = \frac{1}{\alpha_\nu} = \frac{1}{n\,\sigma_\nu}\,, \tag{1.29}$$

where the last equation is for particles with number density n and cross section σ_ν. We have assumed that α_ν does not vary in space. $\langle s_\nu \rangle$ is called the *mean free path*

1.4 Radiative transfer in cosmology

Photons travelling along a ray over cosmological distances redshift. In addition, when counting the rate at which such photons cross a given area, we should account for cosmological time-dilation. Taking account of both of these effects should convince you that the intensity along a ray is unlikely to be conserved, and Eq. (1.22) will look different when considering such rays. How will the new equation look?

To get some physical intuition, it is useful to redo the calculation that yielded cosmological dimming of the surface brightness of an object (as opposed to surface brightness being independent of distance in a non-cosmological setting) - see section 2.8 in the cosmology notes, and recall that $ds = cdt$ for a light ray. So let's write surface brightness $I_\nu \mathrm{d}\nu \propto (1+z)^{-4}$ as

$$I d\nu = I_e\, d\nu_e (a_e/a)^4\,, \tag{1.30}$$

where $I(\nu, t)\mathrm{d}\nu$ is the observed surface brightness, lower than the emitted surface brightness $I_e\, d\nu_e$ by $1/a^4$. Next we use that the photons also redshift,

so that

$$\nu/\nu_e = a_e/a\,.$$

Therefore $I = I_e(a_e/a)^3 = (\nu/\nu_e)^3$, resulting in $I(\nu,t)\,\nu^{-3}$ constant along a ray (in the absence of interactions with matter, of course). When taking the time derivative of this, we need to account again for the fact that ν itself depends on time,

$$\frac{d\nu}{dt} = -\frac{\nu_e\, a_e\, \dot{a}}{a^2} = -H\nu\,,$$

where $H = \dot{a}/a$ is Hubble's constant. Therefore

$$\begin{aligned}\frac{d}{dt}\nu^{-3}I(\nu,t) &= 3\nu^{-4}H\nu\, I(\nu,t) + \nu^{-3}\frac{\partial I(\nu,t)}{\partial t} + \nu^{-3}\frac{\partial I(\nu,t)}{\partial \nu}(-H\nu) \\ &= \nu^{-3}\left(3H\, I(\nu,t) + \frac{\partial I(\nu,t)}{\partial t} - H\nu\frac{\partial I(\nu,t)}{\partial \nu}\right)\,. \qquad (1.31)\end{aligned}$$

Taking into account interactions with matter, which may change $I(\nu,t)\,\nu^{-3}$ as in the non-cosmological case, yields the equation for RT in a cosmological setting:

$$\frac{\partial I(\nu,t)}{\partial t} - H\nu\frac{\partial I(\nu,t)}{\partial \nu} = -3HI(\nu,t) - \alpha_\nu I(\nu,t) + cj_\nu\,. \qquad (1.32)$$

As before, α_ν describes absorption and j_ν emission by matter along the ray; the speed of light appears due to our change of variables $ds \to cdt$. It may be a good approximation to neglect the redshifting terms ($H\nu\partial I(\nu,t)/\partial\nu$ and $-3H\nu$). For example in simulations of reionisation, the photons being tracked are ionising photons. When the Universe is mostly neutral, the mean free path of such photons is so small that they are effectively absorbed 'locally', *i.e.* close to the emitting source, so that $a \approx a_e$.

1.5 Numerical radiative transfer

Monte-Carlo radiative transfer Intuitively rays are like packets of photons, and we could think of writing a radiative transfer code by casting rays in many directions from each source, and solve the RT equation numerically along each ray. This way we can also account for any effect of the absorption of photons by the intervening matter, for example photo-ionisation and photo-heating (discussed in more detail in the next chapters).

Consider for example tracking a ray made-up of many Hydrogen ionising photons through neutral Hydrogen gas. Some of these photons will ionise Hydrogen atoms, making it easier for the next batch of photons to pass. You can see how we would need to iterate how far photons get - because the mean free path of a photon depends on how ionised the gas is along that ray, but the level of ionisation of the gas depends on how many photons pass through. There are other complications, such as when a Hydrogen atom recombines, it will emit an ionising photon (this is called diffuse radiation), and so now every point in space is a potential source of photons from which rays need casting.

This may seem to make the problem intractable, yet such Monte Carlo schemes are widely used for numerical RT. This is because diffuse radiation can often be neglected, and because rays can be cast randomly so that no assumptions on the geometry of the system need be made - a great advantages of this method. If you're interested in this you may want to look-up for example the CRASH code (Graziani et al., 2013) or the one I have used extensively which is called SPHRAY (Altay et al., 2008) (see also my reverse ray-tracing scheme called URCHIN, Altay & Theuns (2013)).

Ray casting on a mesh Similar to MC methods, this involves casting rays from a source, but calculating the interaction of matter within each cell of a Cartesian mesh (which may or not be adaptively refined, see for example Rijkhorst et al. (2006); Mellema et al. (2006)). A little thought shows the typical problem with this: cells far from the source see very few, if any, rays, whereas cells close to the source are probed by numerous rays. This inefficiency is the motivation for ray-splitting algorithms, see *e.g.* Abel & Wandelt (2002).

Moments of the radiative transfer equations The methods discussed above have two main limitations: (*i*) the number of rays that need to be cast increases with the number of sources, and (*ii*) the time-step over which the radiation field is evolved is limited by the speed of light. Moment methods are approximate schemes that do not suffer from these limitations, see for example Gnedin & Abel (2001) and Rosdahl et al. (2013).

Finally it is worth mentioning the scheme based on GADGET described by Pawlik & Schaye (2008), and CLOUDY (Ferland et al., 2013), a widely used detailed RT scheme that includes basically all known interactions between

matter and radiation. We will use CLOUDY in the workshops.

Exercises

1. Derive the relation between optical depth and the change in apparent magnitude of a source.

2. The ISM contains dust grains. Assume spherical grains with radius r_d, and mass density ρ_d. The abundance of grains is described in terms of the dust-to-gas ratio, $\psi = M_d/M_g \approx 10^{-3}$, the ratio of dust mass, M_d, over gas mass, M_g, per unit volume. Find an expression for the optical depth toward the centre of a dusty cloud with radius R and uniform gas density ρ, for given values of r_d and ψ. Evaluate your answer for the case $r_d = 0.1\mu$m, ρ_d=1 g cm^{-3}, $n_{\rm H} = 10^2$ and 10^3 cm^{-3} and $R = 1$ pc. ($n_{\rm H}$ is the hydrogen number density of the gas. Note that ρ_d is the mass density of a single grain.) Would this value apply to IR photons?

3. Free electrons in a plasma scatter photons, with a cross section equal to the Thomson cross section, σ_T. Compute the optical depth due to Thomson scattering to the centre of a cluster of galaxies, using reasonable values of electron density and cluster radius. (Assume a uniform density cluster). Interpret your numerical answer: is Thomson scattering an important process to account for in clusters?

4. Similarly to the previous exercise, compute the Thomson optical depth, as well as the optical depth due to dust, to the nearest star. (At distance $D = 1$ pc, use a typical ISM density of $n_{\rm H}$=1 cm^{-3}, $\psi = 10^{-3}$ and the same dust grain properties as in exercise 2.)

5. The *intergalactic medium* (IGM) at redshifts $z \leq 6$ is observed to be very highly ionised, presumably due to the combined radiation emitted by galaxies and quasars. However we also know that the IGM became (almost completely) neutral following recombination at $z \sim 10^3$: this implies that the IGM was reionised somewhere between $z = 6$ and $z = 10^3$: we currently do not know when this important transition happened. An ionised IGM will Thomson scatter CMB photons, and there is a relation between the reionisation redshift z_r and the Thomson optical depth to the CMB, τ_r. Assume a uniform Universe with given

baryon fraction Ω_b in units of the critical density Ω_c. Derive the relation between τ_r and z_r and compute τ_r if $z_r = 9$ for an Einstein-de Sitter Universe. Use $\sigma_T = 6.65250 \times 10^{-25}$ cm^2.

Chapter 2

Chemical reactions and their application in the ISM

- General properties of reaction rates, form and dimension of reaction coefficients
- Photo-ionisation and recombination
- Interstellar grains: composition of, formation of, extinction due to, radiation pressure on, thermodynamics of

2.1 General properties of reaction rates

The rate at which a chemical reaction such as

$$\mathrm{A} + \mathrm{B} \rightarrow \mathrm{C} \tag{2.1}$$

occurs (for example the formation of CO from O and C) will depend on the properties of the actual system (its temperature, density, abundance distribution), but also on the physics of how likely A and B form C when an interaction does occur. We can separate the rate of such 'collisions' (which will depend on the physical properties of the system) and the cross section of formation (which depends on the physics of the interaction) by writing the production rate of C as

$$\frac{\mathrm{d}n_{\mathrm{C}}}{\mathrm{d}t} = \sigma\, n_{\mathrm{A}}\, n_{\mathrm{B}}\, \langle v_{\mathrm{AB}} \rangle \tag{2.2}$$

$$\equiv k\, n_{\mathrm{A}}\, n_{\mathrm{B}}\,, \tag{2.3}$$

where $n_{\rm X}$ is the number density of species X, and σ ($[\sigma]$= cm^2 is the cross section. The second line defines the rate coefficient $k = \sigma \langle v_{\rm AB} \rangle$ with units $[k] = {\rm cm}^3\ {\rm s}^{-1}$ which will in general depend on the temperature of the gas. Calculation of k for a specific reaction can be a demanding quantum mechanical calculation, and tables of reaction rates and their temperature dependence are continually updated as more accurate versions become available.

2.1.1 Examples

Three-body collisions
The rate of the three body reaction

$$\mathrm{A} + \mathrm{B} + \mathrm{C} \rightarrow \mathrm{D} \tag{2.4}$$

can be written as

$$\frac{{\rm d}n_{\rm D}}{{\rm d}t} = k_3\, n_{\rm A}\, n_{\rm B}\, n_{\rm C}\,. \tag{2.5}$$

From this it is clear that three body reactions occur at a rate $\propto k_3\, n_{\rm C}/k$ as compared to the two-body reaction. If densities are low enough, three body collisions are much less likely than two-body collisions, and can be neglected.

Photo-ionization
The reaction

$$\mathrm{HI} + h\nu \rightarrow \mathrm{HII} + e\,, \tag{2.6}$$

can be written in terms of the density of photons of frequency ν, $n(\nu)d\nu = u(\nu)d\nu/h\nu$ where u is the energy density from Eq. (1.9). Since the relative speed is now the speed of light, c, the rate due to photons with given frequency can be written using Eq. (2.2) as $\sigma_\nu\, n_{\rm HI}\, u_\nu\, c\, d\nu/h\nu$. Photons with energy below $h\nu_{\rm th} = 13.6$ eV are not sufficiently energetic to ionise HI, hence the net ionisation rate due to all photons with $\nu \geq \nu_{\rm th}$ is

$$\frac{{\rm d}n_{\rm HII}}{{\rm d}t} = n_{\rm HI} \int_{\nu_{\rm th}}^{\infty} \sigma(\nu)\, \frac{u(\nu)c}{h\nu}\, d\nu = n_{\rm HI} \int_{\nu_{\rm th}}^{\infty} \sigma(\nu)\, \frac{4\pi J(\nu)}{h\nu}\, d\nu \equiv n_{\rm HI}\, \Gamma\,. \tag{2.7}$$

Here, $J(\nu)$ is the mean intensity from Eq. (1.10) and the last step defines the *photo-ionisation rate* per hydrogen atom, Γ. ($[\Gamma] = {\rm s}^{-1}$). The remainder of the photon's energy, $h_\nu - h\nu_{\rm th}$, goes mainly in the kinetic energy of the liberated electron. Once shared with the other particles through Coulomb collisions, it represents a heating term - *photo-heating* - which we will discuss

later in the context of H II regions.

This expression is relevant when the energy density is known; in the special case where this is dominated by a single source (a star say), with spectrum $L(\nu)$, we can compute the ionisation rate by computing the flux of photons with given frequency at distance r from the source, $F(\nu) = L(\nu)/4\pi r^2$ (valid when neglecting absorption). A similar reasoning as above then yields

$$\frac{\mathrm{d}n_{\mathrm{HII}}}{\mathrm{d}t} = n_{\mathrm{HI}} \int_{\nu_{\mathrm{th}}}^{\infty} \sigma(\nu)\, \frac{F(\nu)}{h\nu}\, d\nu\,. \tag{2.8}$$

It would be useful for you to consider how this expression would change if absorption were accounted for.

Collisional ionization
A sufficiently energetic electron[1] colliding with HI may lead to an ionisation,

$$\mathrm{HI} + e \rightarrow \mathrm{HII} + 2e\,, \tag{2.9}$$

at a rate

$$\frac{\mathrm{d}n_{\mathrm{HII}}}{\mathrm{d}t} = \Gamma_e\, n_{\mathrm{HI}}\, n_e\,. \tag{2.10}$$

The *collisional ionisation rate* Γ_e is temperature dependent, in particular the incoming electron needs to have a kinetic energy greater than the ionisation energy $\chi = 13.6$ eV of Hydrogen: $(1/2)m_e v^2 \geq \chi$. The remainder of the energy of the electron, $(1/2)m_e v^2 - \chi$, is left as kinetic energy.

Recombination
This is the inverse of photo-ionisation,

$$\mathrm{HII} + e \rightarrow \mathrm{HI} + h\nu\,, \tag{2.11}$$

and occurs at a rate governed by the temperature dependent *recombination coefficient* α,

$$\frac{\mathrm{d}n_{\mathrm{HI}}}{\mathrm{d}t} = \alpha(T)\, n_{\mathrm{HII}}\, n_{\mathrm{e}}\,, \tag{2.12}$$

[1] One usually neglects collisions with other particles, for example with protons. Why is this mostly a good approximation?

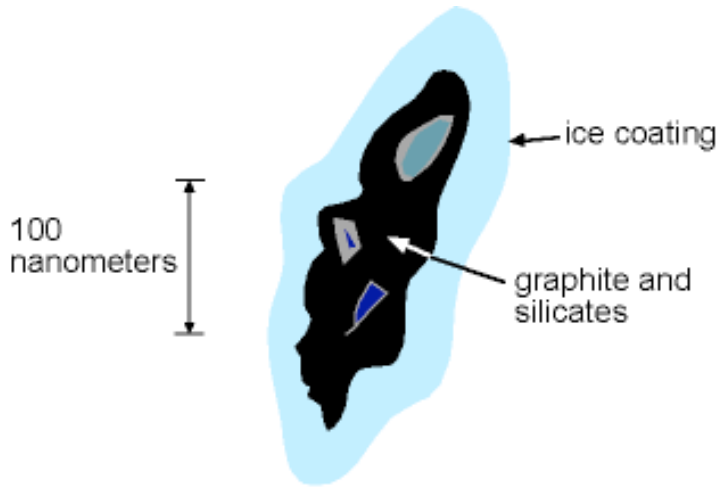

Figure 2.1: Sketch of an interstellar dust grain, from http://www.astro.virginia.edu/class/oconnell/astr121/guide11.html

where T refers to temperature. In ionisation equilibrium, $\mathrm{d}n_{\rm HI}/\mathrm{d}t = 0$, and the three reactions above balance[2]: the equilibrium neutral fraction $x = n_{\rm HI}/(n_{\rm HI} + n_{\rm HII})$ then depends on Γ, Γ_e and α.

You should rightly wonder how the functions $\alpha(T)$, Σ_e, $\sigma(\nu)\cdots$ are related to the quantum mechanics of the atom (Hydrogen, for example). We will discuss this in Chapter 3 below.

2.2 Grains

Grains are typically small (radii $a \sim 0.1 - 1\mu$m; Fig.2.1) solid particles predominantly composed of Si and C, that are important because they cool gas (allowing stars to form), and act as catalysts for chemical reactions (for example the formation of CO molecules). They also are the building blocks for more massive particles, ranging from rocks to eventually planets. Radiation pressure on dust grains is thought to have important effects, from driving

[2]We have neglected the photon produced in the recombination. This is not always a good approximation and we will do better below.

winds in stars to quasars.

2.2.1 Grain formation

A spherical grain (radius a) moving with speed v through the ISM (particle density n, particle mass m_p) grows in radius and mass if these particles stick to the grain, at a rate

$$\frac{\mathrm{d}m}{\mathrm{d}t} = m_p \, (\pi a^2) \, v \, n \, \eta \, , \tag{2.13}$$

where η is a dimensionless *sticking coefficient* that describes the fraction of particles that hit the grain that actually stick to it. Assuming the grain has average density $\rho = m/(4\pi a^3/3)$ then yields

$$\frac{\mathrm{d}a}{\mathrm{d}t} = \frac{\eta \, n \, m_p \, v}{4\rho} \, . \tag{2.14}$$

For typical values in a cold ISM cloud, $a = 10^{-5}$ cm, $n = 10$ cm^{-3}, $v = 10^4$ cm s^{-1} and pure hydrogen gas, it takes $\sim 10^7$ years to build a grain. This is so slow that grain formation in the ISM is not likely to be the dominant formation mechanism of dust grains, and grains are presumably made in denser environments, for example stellar atmospheres (with AGB stars prime candidates). We also know that SNe produce dust (dust has been observed in SN 1987a, Matsuura et al. 2011) which might explain why dust is observed (although not in large quantities) in very high redshift galaxies ($z > 7$) that are so young that the AGB channel for dust production has not had time to contribute significantly to the observed dust.

2.2.2 Extinction by grains

The photon absorption coefficient due to spherical grains of radius a and number density n_d is (see Eq. 1.21)

$$\alpha_\nu = \frac{\mathrm{d}\tau}{\mathrm{d}s} = n_d \, (\pi a^2) \, Q_\nu \, , \tag{2.15}$$

where the *extinction efficiency* Q_ν describes the wave-length dependence of the radiation-dust interaction, $Q_\nu \sim 1$ when $\lambda \ll a$, $Q_\nu \ll 1$ when $\lambda \gg a$.

The amount of dust is often characterised by the dust-to-gas ratio

$$\psi \equiv \frac{M_d}{M_g}\,. \tag{2.16}$$

For gas with mean molecular weight μ this implies for the ratio of number densities of dust particles (n_d) over hydrogen atoms (n_H)

$$\frac{n_d}{\mathrm{n_H}} = \frac{\psi \mu\, m_H}{m_d}\,, \tag{2.17}$$

where m_H is the proton mass. The extinction due to dust , $A_\nu = -2.5 \log{(I_o/I_e)}$, where I_e is the emitted and I_o the observed intensity after absorption (cfr Exercise 1.1), to the centre of a cloud of radius R, is then

$$\begin{aligned} A_\nu &= (2.5 \log e)\, (\pi a^2)\, Q_\nu\, \frac{n_d}{n_H} \int_0^R n_H\, dl & (2.18)\\ &= (2.5 \log e)\, \tau_\nu & (2.19) \end{aligned}$$

where the integral $\int_0^R n_H\, dl \equiv N_H$ is the hydrogen *column density*, and the optical depth

$$\tau_\nu = (\pi a^2)\, Q_\nu\, (n_d/n_H)\, N_H\,. \tag{2.20}$$

The wave-length dependence of Q can often be approximated as

$$Q_\lambda = Q_0 \left(\frac{\lambda_0}{\lambda}\right)^\alpha\,, \tag{2.21}$$

with $\alpha \approx 1{-}2$. Since $\alpha > 0$ absorption by dust is less in the K-band versus the V-band say, and the presence of dust makes a star in the centre of the cloud redder by an amount $A_K - A_V = (Q_K - Q_V)\, (\pi a^2)\, (n_d/n_H)\, N_H\, (2.5 \log e)$.

When observing an external galaxy, dust in the birth clouds of stars, and in the general ISM, both obscure star light so that observed colours differ from emitted colours. Because blue light is absorbed more than red light (think of the colour of the Sun at sunrise/set), this leads to *reddening*. There are (at least) two major complications to unravel this, *i.e.* infer the emitted light from the detected light: (*i*) the amount of dust and the detailed wavelength dependence of Q_ν are mostly unknown, and (*ii*) the dust is not uniformly distributed. In particular, young massive blue stars are mostly embedded in their natal H II regions that are thought to contain much more

dust than the general ISM. Therefore the wavelength dependence of the net extinction curve (ratio of observed over emitted flux) differs substantially from that of the attenuation curve (Q_ν), see e.g. Trayford et al. (2017) who used the SKIRT code to post-process EAGLE galaxies.

The absorbed light heats the dust grains, who then re-emit light in much longer wavelengths, typically the IR. However the re-emitted light is not simply a black-body component, because not all dust grains have the same temperature - this component is often modelled as a 'grey' body. Because a significant fraction of the light is re-radiated, it becomes possible to detect vigorously star-forming galaxies in the IR, and a large fraction of Durham astronomy works on these 'sub-mm' galaxies.

2.2.3 Radiation pressure on grains

Photons absorbed by a grain also transfer their momentum $h\nu/c$ to the dust grain, and this momentum is passed through the gas through gas-dust collisions. This results in a net force acting on the gas. For the grain parameters given as above, the ratio of radiation force over gravitational force, due to a star of luminosity L and mass $M_\star$, at distance R, is

$$\Gamma \equiv \frac{F_{\rm rad}}{F_{\rm grav}} = \frac{(\pi a^2)\, Q\, (L/4\pi c\, R^2)}{G m_d\, M_\star/R^2} = \frac{(\pi a^2) Q\, L}{4\pi\, c\, G\, M_\star\, m_d}\,, \qquad (2.22)$$

an expression similar to that of the Eddington luminosity (L2 stars lectures). Notice that this ratio is independent of R. Typically $\Gamma \gg 1$ in the envelopes of stars, and the dust experiences a strong outward force. The accelerating grains collide with gas particles and drag them along: the presence of dust can be a big factor in driving mass-loss from AGB stars.

The IR photon emitted by the dust grain may scatter several times off other dust grains before it finally escapes from the galaxy. Every time it scatter, is will transfer some of its momentum to the dust. Therefore the effective value of Γ may be higher by a factor of several from single scattering expression above.

2.2.4 Grain thermodynamics

Kirchoff's law

See Kwok, §10.3

A grain being heated by absorbing radiation will re-radiated its heat in the

IR. The relation between its absorption and emission properties follows from considering what happens when placing the grain inside a cavity containing Black Body radiation of temperature T: after reaching thermal equilibrium, the presence of the grain should not affect the BB radiation.

The equation for radiative transfer, Eq. (1.22), with $dI_\nu/ds = 0$ for the thermal equilibrium case, then leads to *Kirchoff's equation* relating absorption and emission[3]:

$$j_\nu = \alpha_\nu \, B_\nu(T_d) \,, \tag{2.23}$$

which describes the emissivity j_ν of the grain in terms of its absorption cross section, α_ν, and the Plank (Black Body) function $B_\nu(T)$ at the dust's temperature T_d. Using Eq. (2.15) to describe the frequency dependence of the grain's absorption cross section, Kirchoff's equation applied to dust grains with number density n_d at temperature T_d, becomes

$$j_\nu = \alpha_\nu \, B_\nu(T_d) = n_d \, (\pi a^2) \, Q_\nu \, B_\nu(T_d) \,. \tag{2.24}$$

Some dust grains may be so small that their temperature is affected significantly by absorbing or emitting a single photon. Such small grains are never in thermal equilibrium, constantly being heated by absorption, and cooling by emission. Detailed models of dust processing take such non-equilibrium effects into account, see e.g. Camps et al. (2016) who computed resulting IR fluxes of EAGLE galaxies.

Dust emission from a uniform cloud

The flux detected at distance D from a uniform cloud with radius R, due to thermal emission from its dust, follows from Eq. (2.24).

$$F_\nu(D) = \frac{4\pi j_\nu \, (4\pi R^3/3)}{4\pi D^2} \tag{2.25}$$

$$= \frac{\psi M_{\rm gas} \, Q_\nu \, B_\nu(T_d)}{(4/3) \, a \, \rho_d \, D^2} \,, \tag{2.26}$$

where ψ is the dust-to-gas ratio, $M_{\rm gas}$ the gas mass of the cloud, and $\rho_d = m_d/(4\pi a^3/3)$ the density of an individual dust grain.

[3]See also *Rybicky & Lightman, §1.5.*

Heating and cooling of dust

Dust placed in an ambient radiation field will absorb radiation and hence be heated, at a rate H per unit volume of

$$\mathrm{H} = \int_\nu \alpha_\nu \, (4\pi J_\nu) \, d\nu \, . \tag{2.27}$$

In the particular case where the radiation field is due to the flux of a single star (radius $R_\star$, effective temperature $T_\star$), at distance D, this becomes (cfr. Eq.1.14)

$$\begin{aligned} \mathrm{H} &= \int_\nu \alpha_\nu \, \frac{\pi B_\nu(T_\star) R_\star^2}{D^2} \, d\nu & (2.28) \\ &= \int \alpha_\nu \, \frac{L_\star(\nu)}{4\pi D^2} \, d\nu \, . & (2.29) \end{aligned}$$

The dust will emit thermal radiation in the IR, which may escape from the system, allowing the system of gas and dust to cool at a rate per unit volume according to Eq. (2.24) of

$$\mathbf{L} = \int (4\pi j_\nu) \, d\nu = \int 4\pi \, n_d \, (\pi a^2) \, Q_\nu \, B_\nu(T_d) \, d\nu. \tag{2.30}$$

In equilibrium, H=L, and

$$\int Q_\nu \, B_\nu(T_d) \, d\nu = \int Q_\nu \, \frac{B_\nu(T_\star)}{4} \, (\frac{R_\star}{D})^2 \, d\nu \, . \tag{2.31}$$

Making the reasonable assumption that the star radiates most of its energy in the visible where $Q_\nu \sim 1$, this simplifies to

$$\int Q_\nu \, B_\nu(T_d) \, d\nu = \frac{1}{4} \, (\frac{R_\star}{D})^2 \int B_\nu(T_\star) \, d\nu = \frac{1}{4} \, (\frac{R_\star}{D})^2 \frac{\sigma T_\star^4}{\pi} \, , \tag{2.32}$$

where σ is the Stephan-Boltzmann constant. In the special case where $Q_\nu = 1$ also for the grain this yields

$$T_d(D) = \sqrt{\frac{R_\star}{2D}} \, T_\star \, , \tag{2.33}$$

and the dust temperature drops with distance from the star $\propto D^{-1/2}$.

We can infer from this reasoning that the dust temperature of a galaxy will depend on its compactness: for the same star formation rate and dust properties, a more compact galaxy (for example at higher redshift) will have *hotter* dust.

2.2.5 Grain chemistry

IR spectroscopy show that most grains are composed of C and Si, often surrounded by an icy mantle (Fig.2.1). Grains are important catalysts for chemical reactions, for example for the formation of molecular hydrogen, where two H atoms adsorbed on a grain, can 'hop' or tunnel across the grain to form H_2, which then may be return to the gas phase (see the cartoon of Fig.2.2). Other molecules may also form in this way.

The rate of molecule formation can be written as

$$\frac{\mathrm{d}n(\mathrm{H_2})}{\mathrm{d}t} = \frac{1}{2}\,\epsilon\, n(\mathrm{H})\, n_d\, (\pi a^2)\, v_{\mathrm{H}}\,, \tag{2.34}$$

where v_{H} is the relative velocity H-grain, and ϵ a dimensionless efficiency factor. The factor 1/2 is because two Hs need to be adsorbed to form H_2. The molecular hydrogen can be destroyed by photons or collisions[4], and together these processes determine the equilibrium molecular fraction.

2.3 Application

A significant amount of all star light produced in the Universe gets re-processed by dust in their host galaxy, and converted from optical and UV to IR wave lengths, see Fig.2.3. As far as the 'spectrum' of the Universe is concerned, there are two other major components: AGN - dominating the high-energy part (X- and γ-rays), and the CMB. This neatly illustrates the three reasons the Universe is not dark: accretion onto black holes (X-ray part of the spectrum), fusion in stars (visible, and through dust re-emission IR part of the spectrum), and matter-anti-matter annihilation, the ultimate origin of the CMB as we discussed in the cosmology lectures.

2.4 Exercises

1. For given values of the photo- and collisional ionisation rates, Γ and Γ_e, and recombination coefficient α, compute the equilibrium neutral fraction, $x = n_{\mathrm{HI}}/(n_{\mathrm{HI}} + n_{\mathrm{HII}})$ in a pure hydrogen gas with number density $n = n_{\mathrm{HI}} + n_{\mathrm{HII}}$. Examine the limits at high- and low-density.

[4]Write-down the rate equations for this.

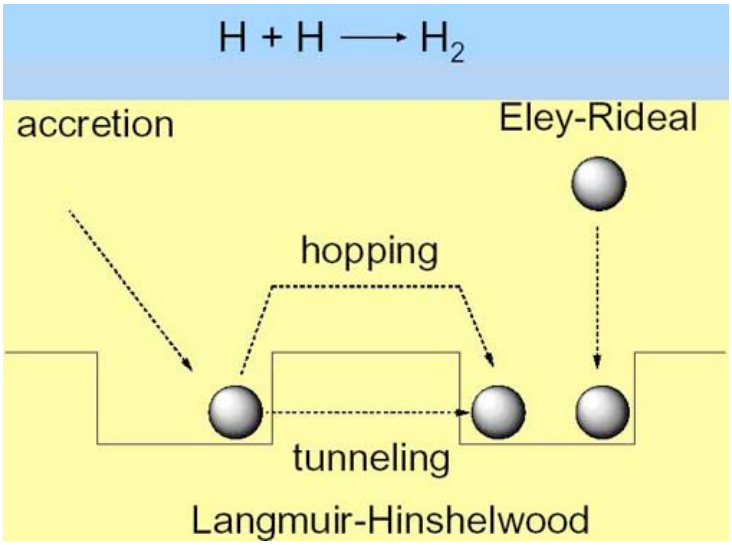

Figure 2.2: The formation of molecular hydrogen on the surface of a dust grain, through adsorption and subsequent tunnelling or hopping of the H-atom over the surface of the grain. Taken from http://www.ifa.hawaii.edu/UHNAI/article3.htm.

2. A pure hydrogen gas with number density $n = n_{\rm HI} + n_{\rm HII}$ is fully ionised, $x = n_{\rm HII}/n = 1$, at time $t = 0$. Assume $\Gamma = \Gamma_e = 0$. Calculate $x(t)$. The recombination time $t_r \equiv 1/\alpha n$. Compute the ionised fraction $x(t_r)$.

3. In a molecular cloud, molecular hydrogen forms on grains and is destroyed by interaction with sufficiently energetic photons. Write the reaction rates for these reactions, and specify the units of all coefficients introduced. Compute the equilibrium molecular hydrogen fraction.

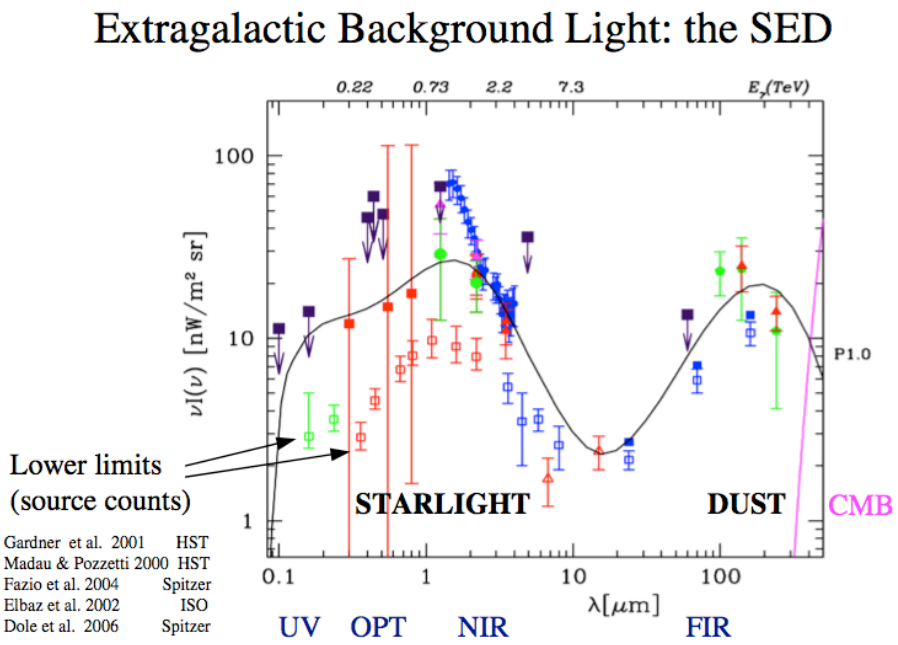

Figure 2.3: The extra-galactic background radiation from the UV-optical to the far IR, from Costamante 2007. At even higher energies the X and gamma-ray background is produced by thermal bremsstrahlung of hot gas and more importantly by the accretion onto black holes, the IR back ground off the right of the plot is due to CMB photons left over from the Big Bang. The UV-optical photons are mostly from star light, but the FIR peak which contains almost as much energy as the UV+optical is mostly star light reprocessed by dust.

Chapter 3

Interaction of radiation with matter

- Review of atomic structure, electron configurations, Hund rules (*Rybicky & Lightman*, §9.1-9.4)
- Radiative transitions, dipole radiation, Einstein coefficients, selection rules, transition rates (*Rybicky & Lightman*, §10.1-10.5)

3.1 Overview of electronic structure of atoms and ions

3.1.1 Schrödinger's equation

The Hamiltonian $\mathcal{H}$ for a multi-electron atom, with N electrons around a nucleus of charge Ze, is the sum of the kinetic energies of all electrons, the attractive Coulomb interaction between each electron and the nucleus, and the repulsive Coulomb interaction between each electron pair:

$$\mathcal{H} = \sum_{i=1}^{N} \frac{p_i^2}{2m} + \sum_{i=1}^{N} \frac{Ze^2}{r_i} - \sum_{i>j} \frac{e^2}{r_{ij}} \,. \tag{3.1}$$

This neglects the kinetic energy of the protons, since we describe the system in the centre of mass of the atom - which, given the large value of the ratio of proton to electron mass, m_p/m_e, nearly corresponds to the system where

the nucleus is at rest. Since $\mathcal{H}$ has no explicit time dependence, the solution to the time-dependent Schrödinger equation, $i\hbar\dot{\psi} = \mathcal{H}\psi$ can be written as $\psi(\mathbf{r}, t) = \psi(\mathbf{r}) \exp(-iEt/\hbar)$ where the spatial part of the wave-function are eigenfunctions of the Hamiltonian, with the energy E the corresponding eigenvalue:

$$\mathcal{H}\psi = E\,\psi\,. \tag{3.2}$$

Eigenfunction are characterised with a set of quantum numbers, denoted n, l, m, for a single-electron system. The operator $\mathcal{H}$ follows from the expression Eq. (3.1) by the transformation $p \to -i\hbar\nabla$. This 'classical' equation misses the spin properties of the electrons (and of the nucleus), which only follow from the relativistic (Dirac) equation. To account for spin, we will simply write the total wave-function as a product of the solution to Eq. (3.2) times the spin state of the electron.

The hydrogen atom

The solution for a hydrogen-like atom with $N = 1$ are analytic, and in spherical coordinates given by

$$\psi(\mathbf{r}) = r^{-1}\, R_{nl}\, Y_{lm}(\theta, \phi)\,, \tag{3.3}$$

where the R_{nl} are associated Laguerre polynomials, and Y_{lm} are spherical harmonics. The quantum numbers n, l and m characterise the energy, total angular momentum, and angular momentum along the z-axis:

$$E_n = -\frac{Z^2 e^2}{2a_0 n^2} = 27.2\,\frac{Z^2}{2n^2}\,\mathrm{eV} \tag{3.4}$$

$$L^2\psi = l(l+1)\hbar^2\psi \tag{3.5}$$

$$L_z\psi = m\hbar\psi\,. \tag{3.6}$$

The angular momentum operator follows from the expression for $\mathbf{L} = \mathbf{r} \times \mathbf{p}$ by making the usual substitution $\mathbf{p} \to -i\hbar\nabla$. In the expression above, $l = 0, 1, \cdots, n-1$ and $m = -l, -l+1, \cdots, l$. These wave-functions are called orbitals, and traditionally[1] $l = 0, 1, 2, 3, \cdots$ are denoted as $s, p, d, f, \cdots$. The energy of an electron only depends on n (degeneracy) with $2n^2$ states for given n, since there are $2(2l+1)$ allowed states[2] for given l.

[1] Originally a spectroscopic notation.

[2] Since $m = -l \cdots l$, there are $2l+1$ valid values of m for a given l, times 2 for the two allowed spin states.

Many-electron systems: LS coupling and Hund's rules

An *Ansatz* for the orbital structure of a many electron system can be made by writing them as linear combinations of the single electron solution[3], for example $\psi(1,2) = \{\psi_{nlm}(1)\psi_{n'l'm'}(2) - \psi_{nlm}(2)\psi_{n'l'm'}(1)\}/\sqrt{2}$, where the particular linear combination expresses the fact that electrons are indistinguishable[4]. The electron configuration is then specified by giving all nlm for each electron. For example the ground state of oxygen (8 electrons) is $1s^2 2s^2 2p^4$.

This *Ansatz* is not an exact solution to the Schrödinger equation, and the coupling induced by the electron-electron interactions cause the ls, m's and s's of individual electrons not to be good quantum numbers anymore (equivalently, the L^2 operator of a *single* electron no longer commutes with $\mathcal{H}$, however the *total* angular and spin momenta are still good quantum numbers (in the sense that states of given energy also correspond to a unique the total angular momentum). The total angular momentum $\mathbf{L} = \sum_j \mathbf{L_J}$ is the vector sum of the individual momenta, and for example for a two-electron system has allowed values $|L_1 - L_2| \leq L \leq L_1 + L_2$ according to the addition rules for angular momentum in quantum mechanics.

The electron structure is then described by 'terms' characterised by the particular values of L and S in what is called LS *coupling* or *Russell-Saunders* coupling. Filled shells have $L = S = 0$ and can be ignored in the coupling, so we only need to specify L and S for the outer most electrons (when in the ground state). Rotational symmetry guarantees that the energies do not depend on m_L or m_S (unless an external field breaks the symmetry).

The spin-interactions in a state with high S will make the electrons stay away from each-other further as compared to a low S state, hence making their Coulomb repulsion lower: this state will hence be *more tightly bound*. Similarly for given S, in a state with higher L, orbiting electrons will be on average further apart from each-other than for lower L, similarly leading to a lower (more bound) energy. This is encapsulated in *Hund's rules*

1. terms with larger spin S tend to lie lower in energy: they are more bound

[3]Recall from Sturm-Liouville theory that these form an orthogonal set of basis functions.

[4]Exchanging $1 \leftrightarrow 2$ changes $\psi \to -\psi$.

2. for a configuration with given S, terms with larger L tend to lie lower in energy

Spin-orbit coupling further breaks the degeneracy of states with the same S and L depending on the total angular momentum $\mathbf{J} = \mathbf{L} + \mathbf{S}$.

In this approximation the electronic state of a system is described by n, S, L and J, and written in the notation

$$^{2S+1}L_J\,,$$

where as before the total orbital angular momentum $L = 0, 1, 2, \cdots$ is written as S, P, D, $\cdots$. An example is shown in Fig.3.1 for two equivalent p electrons. An example is carbon in the ground state, with configuration[5] $1s^2 2s^2 2p^2$: the outer 2 p electron are 'equivalent' and have $n = 2$. Finding which states are allowed is not entirely trivial, see for example wikipedia or Rybicki & Lightman. Once determined, Hund's rules can be used to order them in energy, with L-S coupling breaking the energy degeneracy.

3.2 Dipole radiation

3.2.1 Transition probability

The semi-classical description of the interaction of radiation with atoms treats the atom quantum mechanically but the radiation classically. The Hamiltonian of the system is $\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_1$, where $\mathcal{H}_0$ is the atomic Hamiltonian discussed in the previous section, and the electromagnetic Hamiltonian $\mathcal{H}_1$ is

$$\mathcal{H}_1 = -\frac{e}{hc}\sum_j \mathbf{A}\cdot\mathbf{p}_j\,, \tag{3.7}$$

where $\mathbf{A}(\mathbf{r}, t) = \mathbf{A}(t)\exp(i\mathbf{k}\cdot\mathbf{r})$ is the vector potential, and $\mathbf{p}_j$ is the momentum of each electron.

Treating $\mathcal{H}_1$ as a perturbation on the (initial, i) state of the electrons $|\psi >$ (which are solutions $\mathcal{H}_0|\psi >= E|\psi >$), the probability per unit time,

[5]Since we don't usaully care about filled shells, and because full shells correspond to noble gases, a configuration is of often written as the full shells configuration of the noble atom, plus the remainder. For carbon in the ground state, this would be [He]$2s^2 2p^2$.

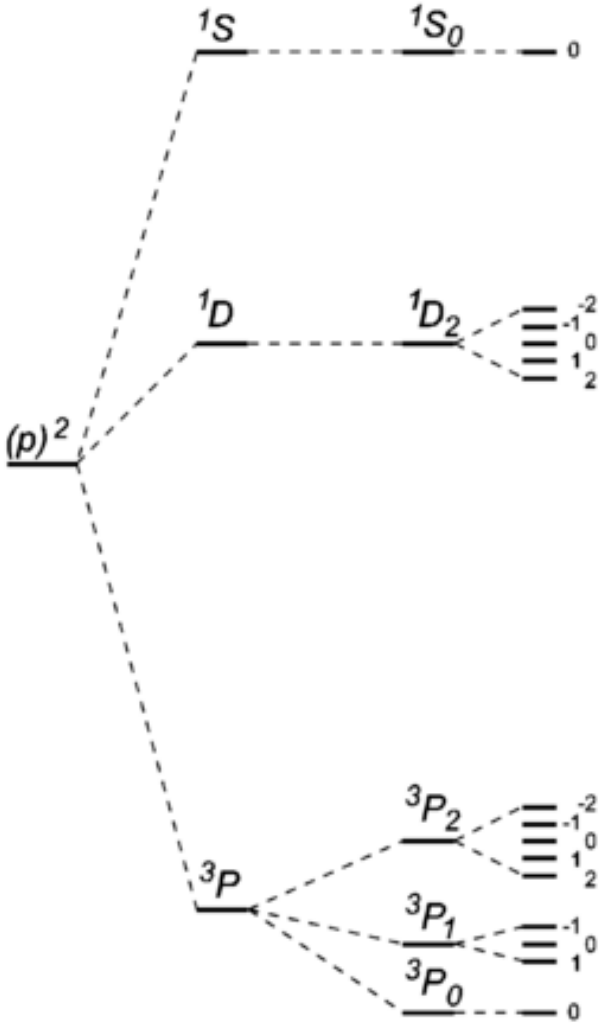

Figure 3.1: Energy levels for two equivalent p-electrons from here; the notation is $^{2S+1}L_J$. The energy degeneracy is broken from left to right due to the following mechanisms. The spin-spin correlation energy makes the singlet states ($S = 0$ hence $2S + 1 = 1$) at higher energy than the triplet state ($S = 1$, hence $2S + 1 = 3$), following Hund's first rule. As a consequence, the states ^1S and ^1D lie above (are less bound) compared to the ^3P state. At given S, states with higher L are at lower energy (so ^1D is lower then ^1S, Hund's second rule). Spin-orbit coupling further breaks the degeneracy of each term according to its J value. The final column to the right shows the remaining $2J + 1$ degenerate levels.

$\mathcal{R}_{if}$, that the radiation will induce a transition from an electronic state $|\psi_i>$ to a state $|\psi_f>$ (final, f), is

$$\mathcal{R}_{if} \propto |<\psi_f|\mathcal{H}_1|\psi_i>|^2 . \tag{3.8}$$

The spatial dependence $\exp(i\mathbf{k}\cdot\mathbf{r})$ of the vector potential can be expanded in a Taylor series,

$$\mathbf{A}(\mathbf{r}) = \mathbf{A}_0 \exp(i\mathbf{k}\cdot\mathbf{r}) \approx \mathbf{A}_0 \left(1 + i\mathbf{k}\cdot\mathbf{r} + \cdots\right), \tag{3.9}$$

and the *dipole approximation* limits this expansion to the first term (*i.e.* $\mathbf{A}_0$). Higher-order contributions are typically smaller by factors v/c and can usually be neglected, except if the dominant dipole term is zero from symmetry arguments. Such 'forbidden' lines occur then because of these higher-order terms, and correspond to *electric quadrupole*, *magnetic dipole* and other higher-order terms. We will see below that such forbidden lines can actually be very strong in the ISM or in stellar spectra, for example forbidden [OIII] transitions (forbidden lines are denoted by the square brackets []) are the strongest lines observed in planetary nebulae.

In the dipole approximation the transition probability is then given by the volume integral

$$\mathcal{R}_{if} \propto |\int \psi_f^\dagger \, \mathbf{A}_0 \cdot \sum_j \mathbf{p}_j \, \psi_i \, dV|^2 . \tag{3.10}$$

By applying the commutation relation[6] $[\mathbf{r}_j, \mathcal{H}_0] = i\hbar\mathbf{p_j}$, this can be written as

$$\mathcal{R}_{if} \propto |\hbar^{-1} \int \psi_f^\dagger \left(\mathbf{A}_0 \cdot \left\{ \sum_j (\mathbf{r_j}\mathcal{H}_0 - \mathcal{H}_0\mathbf{r_j}) \right\} \right) \psi_i \, dV|^2 \tag{3.11}$$

$$\propto |\hbar^{-1}(E_i - E_f) \int \psi_f^\dagger \, (\mathbf{A}_0 \cdot \sum_j \mathbf{r_j}) \, \psi_i \, dV|^2 \tag{3.12}$$

since $\mathcal{H}_0|\psi_{i,f}>= E_{i,f}\,|\psi_{i,f}>$. This expression shows that the transition rate is proportional to the (square) of the expectation value of the *electric dipole operator*, $\mathbf{d} = e\sum_j \mathbf{r}_j$, hence the name *dipole radiation*.

[6]Demonstrate this as an exercise.

3.2.2 Selection rules

The transition probability Eq (3.12) is zero and hence not 'dipole allowed' if the initial and final states do not adhere to certain selection rules.

1. *Laporte's rule*: Since the dipole operator is odd, the parity of initial and final state should be opposite, since otherwise the integral and hence the probability is identically zero

2. The transition probability for a many-electron system is a sum of terms such as

$$\int \psi_a^\dagger(1)\,\psi_b^\dagger(2)\,\psi_c^\dagger(3)\cdots\ \mathbf{r}_j\ \psi_{a'}(1)\,\psi_{b'}(2)\,\psi_{c'}(3)\ \cdots d^3x_1\,d^3x_2\,d^3x_3\cdots \tag{3.13}$$

   where $a, b, c, \cdots$ and $a', b', c', \cdots$ are the electron configurations in the final and initial state, respectively. For an electron $i \neq j$, the integral over all space will be zero unless the initial and final states are the same, since the ψ_i are orthogonal functions. Hence the *one electron jump rule*: all electrons orbitals should be the same, except for one.

3. Since the dipole operator does not involve spin, the spin of the electron can not change.

4. The selection rules

   - $\Delta l = \pm 1$
   - $\Delta m = 0, \pm 1$

   follow from examining the dipole operator for the jumping electron, see *Rybicky & Lightman exercise 10.6*. This basically follows from writing the angular part dipole operator $\mathbf{r}$ itself in Legendre polynomials, and examining the orthogonality of products of Legrende polynomials, see exercise 1 below.

As remarked earlier transitions between states that do not follow these selection rules, although forbidden in the dipole approximation, may nevertheless occur due to higher-order terms that we have neglected here. An example shown in Fig.3.2 are forbidden transitions in OIII (doubly ionised oxygen), for which the two valence electrons have the $2p^2$ configuration discussed earlier.

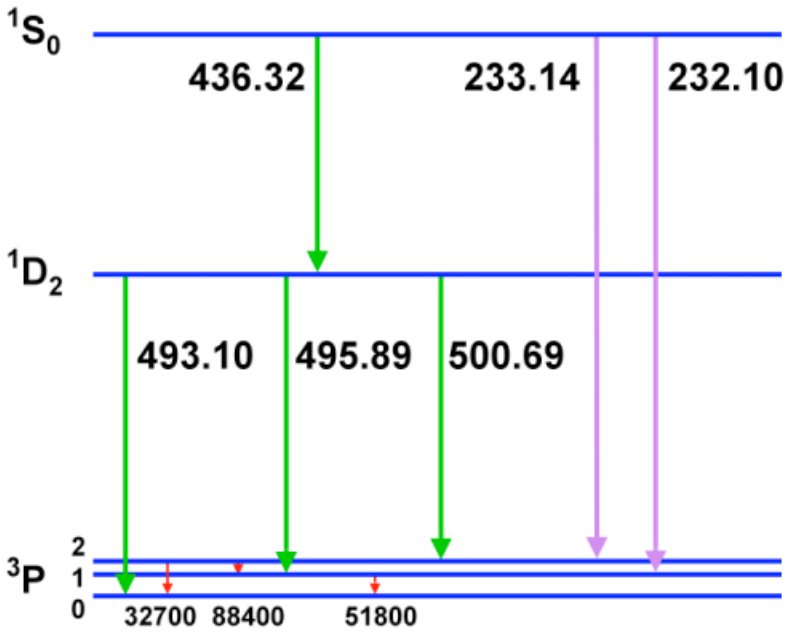

Figure 3.2: The transitions $^1D_2 \rightarrow^3 P_2$ at $\lambda = 500.7$ nm and $^1D_2 \rightarrow^3 P_1$ at 495.9 nm in OIII are dipole-forbidden, yet are usually by far the strongest nebular lines in Planetary Nebulae. These lines given such nebulae their typical green colours, see this example from the SDSS.

3.2.3 Examples

Bound-bound transitions

The dipole operator can be evaluated when the electron configurations of initial and final state are known. These can be computed exactly for the hydrogen atom, and involve integrals of the form

$$|\langle\psi_f|\mathbf{r}|\psi_i\rangle|^2 \propto \int (rR_{nl})\, r\, (rR_{n'l'})\, r^2 dr\,, \tag{3.14}$$

in terms of the radial dependence of the orbitals (Laguerre polynomials R_{nl}).

Bound-free transitions (photo-ionisations)

If the electron in the final state of the interaction is unbound, its wavefunction is that of a free particle, $|\psi_f\rangle \propto \exp(-i\mathbf{q}\cdot\mathbf{r})$. The transition amplitude is then

$$\begin{aligned} |\langle\psi_f|\mathbf{A}_0\cdot\mathbf{p}|\psi_i\rangle &\propto \Big|\int \psi_f^\dagger(\mathbf{A}_0\cdot\nabla)\,\psi_i\, d^3x\Big| && (3.15)\\ &\propto |\mathbf{A}_0\cdot\int \psi_i^\dagger\,\nabla\psi_f\, d^3x| && (3.16)\\ &\propto |(\mathbf{A}_0\cdot\mathbf{q})\int \psi_i^\dagger\,\psi_f\, d^3x|\,, && (3.17)\end{aligned}$$

which depends on the momentum q of the free electron. Photo-ionisation will not take place if the energy $h\nu$ of the incoming electron is less than the binding energy χ of the atom, $h\nu \geq h\nu_{\rm th} = \chi$. Above this threshold the cross section falls approximately $\propto \nu^{-3}$. The net cross section is thus

$$\sigma_{bf} = \begin{cases} 0, & \text{if } h\nu < h\nu_{\rm th} \\ \sigma_0(\nu_{\rm th}/\nu)^3, & \text{if } h\nu \geq h\nu_{\rm th}\,. \end{cases} \tag{3.18}$$

Here, σ_0 is the cross section at the threshold. The presence of such thresholds introduces characteristic *photoelectric absorption edges* in the spectra of sources such as for example AGN whose light is reprocessed by intervening gas; see Fig.3.3 for an example.

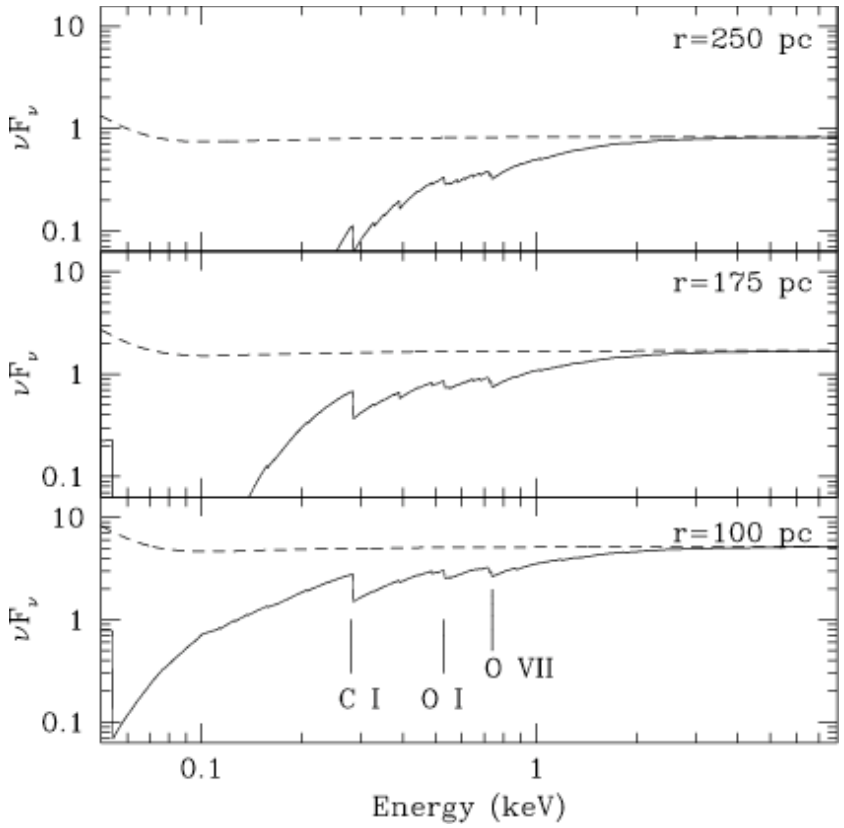

Figure 3.3: Photoelectric absorption edges imprinted by intervening gas in a mock AGN spectrum, from Ballatyne et al, A&A 409, 503. Sufficiently energetic photons can ionise the intervening matter to a higher ionisation state, and this increases the opacity of the intervening medium. Photons need to have an energy above the ionisation energy of the particular ion, $h\nu \geq h\nu_{\rm th}$, therefore the ionisation edge appears as a saw-tooth, with depth σ_0 and recovering to the continuum $\propto (\nu_{\rm th}/\nu)^3$ see Eq. (3.18). Ionic transitions that cause the edges are indicated.

Free-bound transitions (recombinations)

The cross section σ_{fb} for free-bound transitions can be found by considering the *detailed balance* relation for a system of atoms and ions, in ionisation equilibrium with an ambient Black Body radiation field. The (recombination) rate for the reaction

$$N^+ + e \rightarrow N\,, \tag{3.19}$$

in terms of the density of ions (n_+) and electrons (n_e) is

$$\frac{\mathrm{d}n_+}{\mathrm{d}t}|_{\mathrm{rec}} = -n_+\, n_e\, \sigma_{fb}(v) v\, f(v)\, dv\,, \tag{3.20}$$

where $\sigma_{fb}(v)$ is the recombination cross section for electrons moving with velocity v, and $f(v)$ is the fraction of electrons with velocity in the interval $[v, v+dv]$ (this is an application of Eq.2.2). The corresponding photo-ionisation rate is

$$\frac{\mathrm{d}n_+}{\mathrm{d}t}|_{\mathrm{ph}} = n_0\, \frac{4\pi B_\nu}{h\nu}\, [1 - \exp(-h\nu/kT)]\, \sigma_{bf}(\nu)\, d\nu\,, \tag{3.21}$$

see Eq. (2.7), in case the radiation field is that of a Black Body, $J_\nu = B_\nu$. The term $1 - \exp(-h\nu/kT)$ accounts for stimulated emission, and n_0 is the number density of neutrals. In equilibrium the net rate $\mathrm{d}n_+/\mathrm{d}t = 0$, hence

$$n_+\, n_e\, \sigma_{fb} f(v) \frac{h\nu}{m} = n_0\, \exp(-h\nu/k_T)\, \frac{4\pi}{h\nu}\, \frac{2h\nu^3}{c^2}\, \sigma_{bf}\,, \tag{3.22}$$

where we used energy conservation,

$$h\nu = mv^2/2 + \Delta E\,, \tag{3.23}$$

with ΔE the ionisation energy, and Eqs. (A.1) and (A.2) for the BB and Maxwell distributions, applicable in thermal equilibrium. Using Saha's equation, Eq. (A.8), to the ratio $n_+\, n_e/n_0$ then yields the *Milne relation* between photo-ionisation and recombination cross sections:

$$\frac{\sigma_{bf}}{\sigma_{fb}} = \frac{g_e\, g_{n_+}}{2 g_{n_0}}\, \frac{m^2 c^2 v^2}{h^2 \nu^2}\,. \tag{3.24}$$

The g factors represent the statistical weights of each term.

3.3 Exercises

1. Derive the selection rules $\Delta l = \pm 1$, $\Delta m = 0, \pm 1$. Hint: write the angular dependence of orbitals as $P_l^m(\mu)\,\exp(im\phi)$ in terms of the associated Legendre functions, where $\mu = \cos(\theta)$, and apply the recurrence relations

$$(2l+1)\,\mu P_l^m = (l-m+1)\,P_{l+1}^m + (l+m)P_{l-1}^m \quad (3.25)$$

$$(2l+1)\,\sqrt{1-\mu^2}P_l^{m-1} = P_{l-1}^m - P_{l+1}^m\,. \quad (3.26)$$

   Write the dipole operator $\propto \mathbf{r}$ in terms of $((x \pm iy), z)$ or $(\sin(\theta)\exp(\pm i\phi), \cos(\theta))$.

2. C IV, Si IV and O VI all produce characteristic doublets. Show that these ions all have the same ground state electron configuration, which is also the same as that of Alkali metals. Just as in Alkali metals, the doublet results from transitions between the excited (p) state, and the ground (s) states. Write down the state for each of these three configurations. Use Hund's rules to rank them in energy. Are these transitions dipole allowed? Consider the degeneracy levels of the excited states: what is the ratio of line strengths for the two lines in the doublet? [Hint: for given S and L, states with higher J are less bound (Lande's rule).]

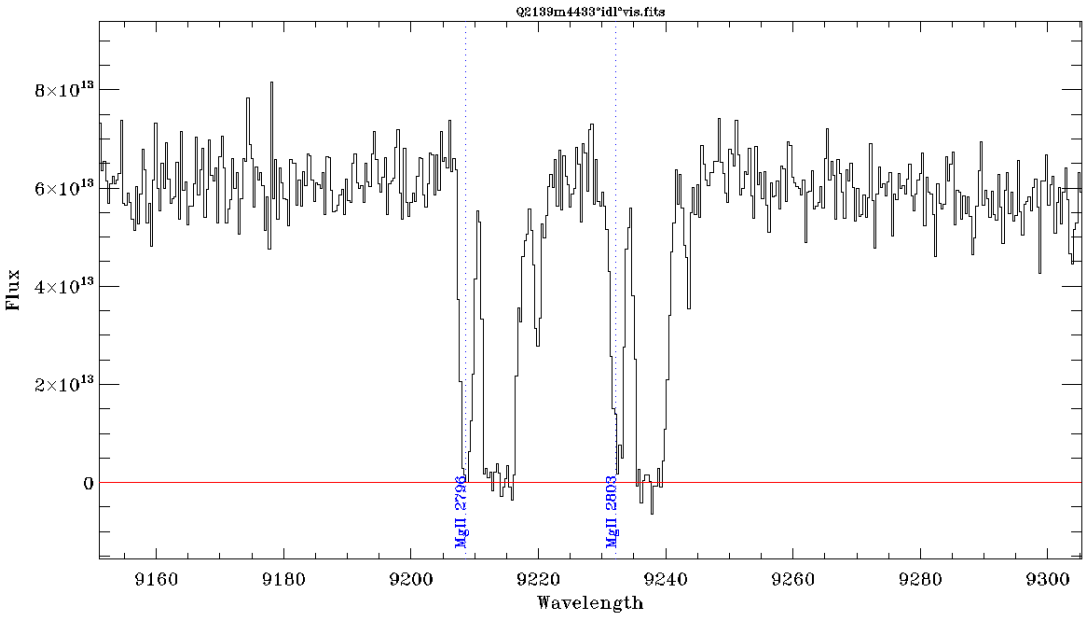

Figure 3.4: Example of a doublet (Mg II, in this case). Notice how each feature on the left (the trio of lines around between 9200 and 9220Å) is repeated on the right (the trio of lines between 9230 and 9250Å). For some components of the doublet it is clear that the shorter wavelength component (higher energy) is stronger than the corresponding weaker component. Why is that no as obvious for the stronger components that reach flux ~ 0?

Chapter 4

Nebulae

Dyson & Williams, 5.1–5.3, Kwok §5.12

- H II regions
  - jump condition
  - on-the-spot approximation, case A and case B recombination
  - ionization level and sharpness of ionised region
  - heating and cooling processes, equilibrium temperature
  - importance of line cooling by metals
- temperature sensitive line ratios

4.1 H II regions

H II regions are interstellar gas clouds in which the hydrogen in part of the cloud is photo-ionised by a hot star, see for example images and explanation here. Usually this signals the end of star formation in the initially molecular cloud, as the photo-ionised gas is also heated and this makes the cloud disperse. The ionising star needs to be massive, since the Black Body spectrum of lower mass stars produces very few photons with energy ≥ 13.6 eV required to ionise H I. The presence of other elements, such as He and trace fractions of metals, affects the ionisation and thermal structure of the cloud: line cooling by metal transitions reduces the temperature of the cloud.

4.1.1 Description in terms of a jump condition

Consider a spherical cloud, initially completely neutral, consisting of hydrogen with uniform number density $n_{\rm HI} = n$. At time $t = 0$, a hot star at the centre of the cloud starts ionising its surroundings, emitting ionising photons at a constant rate $\dot{N}_\gamma$. The gas inside the *ionisation front* at radius R will be mostly ionised, 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

$$4\pi\, R^2\, n_{\rm HI}\, \frac{{\rm d}R}{{\rm d}t} + \alpha\, n_{\rm HII}\, n_e\, \frac{4\pi}{3} R^3 = \dot{N}_\gamma\,, \tag{4.1}$$

where $n_e = n_{\rm HII}$ is the density of electrons and ions, and α is the recombination coefficient. This *jump* conditions expresses the fact that an ionising photon either ionises a neutral atom for the first time increasing R (first term) or compensates for a recombination within the ionised region (second term). *If* we assume the gas inside the ionised region to be completely ionised, then $n_{\rm HII} = n_e = n$ inside the ionized region, and $n_{\rm HI} = n$ in the neutral zone, and Eq. (4.1) can be solved exactly:

$$R(t) = R_S\, (1 - \exp(-t/t_r))^{1/3}\,, \tag{4.2}$$

where the recombination time $t_r = (\alpha\, n)^{-1}$ and R_S is the *Strömgren* radius.

4.1.2 Case B recombination, and on-the-spot approximation

See also *Rybicky & Lightman §10.5*
The general solution above shows that the ionisation front stalls at the Strömgren radius R_S for times $t \gg t_r$. The system then reaches an equilibrium where photo-ionisations balance recombinations. The photo-ionisation rate Γ from Eq. (2.7) is

$$\Gamma = \int_{\nu_{\rm th}}^{\infty} \sigma(\nu)\, \frac{4\pi J(\nu)}{h\nu}\, d\nu\,, \tag{4.3}$$

where the mean intensity in this particular case of a single source with spectrum $L(\nu)$ at distance R is

$$4\pi J_\nu = \frac{L(\nu)}{4\pi R^2}\, \exp(-\tau_\nu)\,. \tag{4.4}$$

This follows from considering the flux received at distance R from a source: the first factor combines the definition of mean intensity $4\pi J_\nu = \int I_\nu \, d\Omega$ (Eq. 1.10) with the equation that relates flux and intensity, $F = \int I \, \cos\theta \, d\Omega$ (Eq. 1.14), the second factor $\exp(-\tau_\nu)$ takes into account absorption. The recombination rate per unit volume is $\alpha \, n_{\rm HII} \, n_e$ (Eq. 2.12). In equilibrium, ionisations balance recombinations, hence

$$n_{\rm HI} \, \Gamma = \alpha \, n_{\rm HII} \, n_e \, , \tag{4.5}$$

note that we have neglected collisional ionizations here. We will see later that the temperature in an H II region is sufficiently low for this to be a good approximation.

The recombinations in the the RHS of Eq. (4.5) involve the capture of a free electron to a given energy level n of the hydrogen atom followed by a cascade of the excited electron to lower energy states until it reaches the ground state $n = 1$. We can compute the probability for capture to level n along the lines described in section 3.2.3, let's denote this recombination coefficient as α_n, a function of temperature T. The recombining gas will hence emit a series of hydrogen emission lines, corresponding to the excited electron making transitions $n \rightarrow n' < n$. The Balmer lines of which Hα has $n = 3 \rightarrow n' = 2$, at a wavelength of $\lambda = 6563 \mathring{A}$, give H II regions their characteristic red colour. The total recombination rate is then simply the sum of recombinations to all energy levels n: this is called the *case A recombination coefficient*: $\alpha_A = \sum_n \alpha_n$.

However consider recombinations directly to the ground state, which occur at a rate governed by α_1. Such a transition will result in the emission of a photon with energy $h\nu \geq 13.6$ eV, since the initially free electron had energy greater than the ionisation energy of H I. The resulting photon therefore has enough energy to ionise (another) H I: how should we treat this so-called *diffuse radiation*[1]?

One way would be to describe ionisations in the H II region as being due either to photons from the source, or photons resulting from the diffuse radiation field. This makes the problem too hard for an analytical description,

[1]Diffuse since the whole ionised part of the H II region acts as a source of ionising photons.

but it is possible to treat it numerically. Another way is to make the following approximation: assume that the photon produced by a recombination to the ground state *ionises a neutral hydrogen at the same position*: this is called the *on-the-spot approximation*. Recombinations to the ground state then have no effect at all, since each recombination is balanced by a new ionisation *at the same location*, not affecting the neutral fraction - or indeed anything else.

The H II region in the on-the-spot approximation is then given by Eq. (4.5), except we should exclude recombinations to the ground state: the recombination coefficient is then the *case B* coefficient $\alpha_B \equiv \alpha_A - \alpha_1 = \sum_{n=2}^{\infty} \alpha_n$: the sum of recombinations to all levels, *excluding* those to the ground state $n = 1$.

The equilibrium relation, Eq. (4.4), can be integrated to the edge of the H II region, at position R_E, to yield

$$\int_0^{R_E} n_{\rm HI}\, 4\pi R^2\, dR \int_{\nu_{\rm th}}^{\infty} \frac{L(\nu)}{h\nu}\, \frac{1}{4\pi R^2}\, \exp(-\tau_\nu)\, \sigma_\nu\, d\nu = \int_0^{R_E} 4\pi\, R^2\, \alpha_B\, n_{\rm HII}\, n_e\, dR\,. \tag{4.6}$$

Interchanging the integrals, the LHS becomes

$$\int_{\nu_{\rm th}}^{\infty} \frac{L(\nu)}{h\nu}\, \left\{ \int_0^{R_E} \exp(-\tau_\nu)\, n_{\rm HI} \sigma_\nu\, dR \right\} d\nu\,. \tag{4.7}$$

Using the relation between neutral density and cross-section, yields that the optical depth $d\tau_\nu = \sigma_\nu\, n_{\rm HI}\, dR$. Therefore the integral in brackets is $\int_0^\infty \exp(-\tau_\nu)\, d\tau_\nu = 1$, hence the LHS becomes $\int_{\nu_{\rm th}}^{\infty} L(\nu)/h\nu\, d\nu = \dot{N}_\gamma$, the rate at which the central source emits ionising photons. The ionisation equilibrium conditions then becomes

$$\int_0^{R_E} 4\pi\, R^2\, \alpha_B\, n_{\rm HII}\, n_e\, dR = \dot{N}_\gamma\,, \tag{4.8}$$

the familiar expression for the Strömgren radius, $R_E \equiv R_S$, and with the case B recombination coefficient.

4.1.3 The ionisation level within the H II region

In ionisation equilibrium, the ionised fraction $x = n_{\rm HII}/n$ is a solution to

$$(1 - x)\Gamma = \alpha_B\, x^2\, n\,. \tag{4.9}$$

At distance R from the star, the ionisation rate is

$$\Gamma = \int_{\nu_{\rm th}} \frac{L(\nu)}{h\nu} \frac{1}{4\pi R^2} \exp(-\tau_\nu)\, \sigma_\nu \, d\nu \tag{4.10}$$

$$\approx \frac{\sigma_{\nu_{\rm th}}}{4\pi R^2} \int_{\nu_{\rm th}}^{\infty} \frac{L(\nu)}{h\nu} \, d\nu \tag{4.11}$$

$$\approx \frac{\sigma_{\nu_{\rm th}} \dot{N}_\gamma}{4\pi R^2} \, . \tag{4.12}$$

The first approximation is to neglect the $\exp(-\tau_\nu)$ by assuming the gas to be very highly ionised, the second approximates the photo-ionisation cross section $\sigma_\nu \approx \sigma_{\nu{\rm th}}$. We can use the following typical parameters to judge how well these approximations work: assume the ionising source is a single O-star which typically emits $\dot{N}_\gamma = 10^{49}$ s^{-1}, whereas the cloud has $n = 10^2$ cm^{-3}, and $\sigma_{\nu_{\rm th}} = 6.3{\times}10^{-18}$ cm^2. At a distance of $R = 1$ pc, $\Gamma \approx 5.3{\times}10^{-5}$ s^{-1}. The case B recombination coefficient $\alpha_B \approx 2.6 \times 10^{-13}$ cm^3 s^{-1} at a temperature of $T = 10^4$ K. Substituting these numbers yields the equation

$$\frac{\Gamma}{\alpha_B \, n} \approx \frac{5 \times 10^{-7}\,{\rm s}^{-1}}{2.6 \times 10^{-13}\,{\rm cm}^3\,{\rm s}^{-1} \times \ 10^2\,{\rm cm}^{-3}} \approx 10^4 \ = \frac{x^2}{1-x} \, , \tag{4.13}$$

hence $x \approx 1 - 10^{-4}$ and the H II region is indeed very highly ionised.

4.1.4 The width of the H ii region

H II regions are very sharp, *i.e.* the distance over which the gas goes from very highly ionised, to almost completely neutral, is small in general. The width can be estimated by solving

$$\frac{x^2}{1-x} = \frac{\Gamma}{\alpha_B \, n} \tag{4.14}$$

$$d\Gamma \approx -\Gamma \, \sigma \, n_{\rm HI} \, dR = -\Gamma\sigma \, (1-x) \, n \, dR \, , \tag{4.15}$$

which combines the equation for ionisation equilibrium, with an approximation of how quickly Γ changes with R close to the edge of the ionisation front[2].

[2]The approximation neglects the geometric $1/R^2$ decrease in Γ.

These equations combine to

$$d\lambda = -\frac{2-x}{x\,(1-x)^2}\,dx\,, \tag{4.16}$$

where $d\lambda = \sigma\, n\, dR$. λ is a dimensionless measure of distance, $d\lambda = dR/\langle s_\nu \rangle$, where $\langle s_\nu \rangle$ is the mean free path from Eq. (1.29). The above differential equation can be solved for $\lambda(x)$, and the solution shows λ diverges only logarithmically for $x \to 0$. For example for $\lambda = 10$, $x \ll 1$, yet for the H II regions parameters from before this corresponds to $\Delta R \approx 10^{-3}$ pc$\ll R_S$: the ionisation fraction x goes from very large $x \sim 1$ to very small $x \ll 1$ over a distance much smaller than the Strömgren radius.

4.1.5 The temperature of H ii regions

Heating

Neutral hydrogen in an H II regions gets photo-ionised by photons with energy $E = h\nu$ larger than the ionisation energy, $h\nu_{\rm th}$ =13.6eV. The excess energy $\Delta E = h\nu - h\nu_{\rm th}$ goes initially (mostly) into kinetic energy of the liberated electron, which shares that energy through collisions. This heats the gas: the process is called *photo-heating*.

The heating rate H per unit volume, follows from multiplying the ionisation rate with the energy ΔE liberated per ionisation,

$$\mathrm{H} = n_{\rm HI} \int_{\nu_{\rm th}}^{\infty} \frac{4\pi J_\nu}{h\nu}\,(h\nu - h\nu_{\rm th})\,\sigma_\nu\,d\nu\,. \tag{4.17}$$

The ionisation rate per hydrogen atom due to photons with frequency ν is $(4\pi J_\nu/h\nu)\,\sigma_\nu$, and this is multiplied by the energy liberated $(h\nu - h\nu_{\rm th})$ per ionisation to get the heating rate. Using Eq. (4.9) for the equilibrium neutral fraction $n_{\rm HI}/n = 1 - x$ yields

$$\mathrm{H} = \frac{x^2 \alpha_B n^2}{\Gamma} \left\{ \int_{\nu_{\rm th}}^{\infty} \frac{4\pi J_\nu}{h\nu}\,(h\nu - h\nu_{\rm th})\,\sigma_\nu\,d\nu \right\} \tag{4.18}$$

$$\approx \alpha_B\, n^2\, \frac{\epsilon}{\Gamma}\,, \tag{4.19}$$

where the last step defines

$$\epsilon \equiv \int_{\nu_{\rm th}}^{\infty} \frac{4\pi J_\nu}{h\nu}\,(h\nu - h\nu_{\rm th})\,\sigma_\nu\,d\nu\,, \tag{4.20}$$

and the approximation $x \approx 1$ was used for the ionised fraction within the H II region. Verify that H has the correct units.

The heating rate H depends on the ratio ϵ/Γ: note that since both are proportional to the intensity J, the heating rate is *independent* of J. At first this seems curious: surely the gas will be heater *more* if the intensity of the ionising radiation, and hence Γ, is increased. However, recall that only the neutrals get heated. Increasing J indeed increases the heating rate ϵ, but at the same time it decreases the neutral fraction: these exactly compensate. Note that H depends on the *shape* of the ionising spectrum (a source with relatively more high energy photons heats the gas more), and also on the density of the gas (higher density gas gets heated more).

An estimate of the heating follows from equating the mean energy per photon to the effective temperature of the ionising star, $\epsilon/\Gamma \approx (3/2)\, kT_\star$.

Cooling

As the ionised gas recombines, the excess energy of the recombining electron is radiated away[3], which presents a loss term for the thermal energy of the gas: this process is called *radiative cooling*. The cooling rate per unit volume, L, can be estimated by multiplying the recombination rate per unit volume, with the energy ΔE lost per recombination, $\Delta E \approx (3/2)\, kT$, with T the temperature of the nebulae:

$$\mathrm{L} \approx \alpha_B\, n_{\mathrm{HII}}\, n_e\, \eta\, kT\,. \tag{4.21}$$

If all electrons had equal chance of recombining, we would expect $\eta = 3/2$, the mean energy per particle. However, low energy electrons have a higher chance of recombining than higher energy electrons and in practise $\eta \approx 1$, slightly smaller than expected.

The equilibrium temperature of a pure hydrogen H ii region

In equilibrium, H = L implies $x^2\, \alpha_B\, n^2\, (3/2) kT_\star = x^2\, \alpha_B\, n^2\, \eta\, kT$, or $T \approx (3/2\eta)\, T_\star > T_\star$. This is independent of distance and density, but depends on

[3]The photon can escape provided the optical depth is small enough.

the stellar temperature. The nebula is hotter than the star, because $\eta < 3/2$. Observed H II regions have lower temperatures than this: this is because particle *collisions*, in particular those involving metals, contribute significantly to the cooling rate. We describe this process next.

Cooling due to metals: collisionally excited line radiation

Osterbrock & Ferland, §3.5
Collisionally excited line cooling occurs when an atom or ion is collisionally excited to an energy level ΔE above the ground state, then returns to its ground state by a radiative transition. When this photon with energy ΔE escapes from the system, it takes away energy, hence this process acts to cool the gas. The rate at which this occurs can be found by first considering the system to be in equilibrium, in which case we can apply the principle of detailed balance. The resulting atomic relations also apply outside of equilibrium[4].

Consider a 2-level atom, with the excited level $n = 2$ at energy ΔE_{12} above the ground state $n = 1$. Let $\sigma_{12}(u)$ be the velocity-dependent cross section for collisional excitations $1 \rightarrow 2$, and similarly for σ_{21}. Detailed balance in equilibrium requires

$$n_e\, n_1\, u_1\, \sigma_{12}(u_1)\, f(u_1)\, du_1 = n_e\, n_2\, u_2\, \sigma_{21}(u_2)\, f(u_2)\, du_2\,, \tag{4.22}$$

an application of Eq. (2.2), where the Maxwellian distribution $f(u)$ at the equilibrium temperature T is given by,

$$f(u) \propto u^2\, \exp(-mu^2/2kT)\,. \tag{4.23}$$

Here, n_1, n_2 and n_e are the number densities of ground state, excited state, and electrons, respectively. In equilibrium, the ratio of occupation numbers follows the Boltzmann distribution (Eq. A.5), evaluated at the *same* equilibrium temperature T:

$$\frac{n_2}{n_1} = \frac{g_2}{g_1}\, \exp(-\Delta E/kT)\,. \tag{4.24}$$

[4]As another example examine the derivation of the Einstein relations in the Appendix.

Energy conservation implies that

$$(1/2)mu_1^2 = (1/2)mu_2^2 + \Delta E_{12}\,. \tag{4.25}$$

Now define the dimensionless energy-dependent cross section, $\Omega(1,2;E)$, as

$$\sigma_{12}(u) \equiv \frac{\pi\hbar^2}{m^2u^2}\,\frac{\Omega(1,2;E)}{g_1}\,. \tag{4.26}$$

Verify this is indeed dimensionless, and that from the previous discussion it follows that

$$\sigma_{21}(u) \equiv \frac{\pi\hbar^2}{m^2u^2}\,\frac{\Omega(1,2;E)}{g_2}\,. \tag{4.27}$$

With this definition, the collisional de-excitation rate is

$$\begin{aligned}
n_e n_2 \int_0^\infty \quad u_2 \quad \sigma_{21}(u_2)\, f(u_2)\, du_2 &= n_e n_2 \int_0^\infty u_2\, \frac{\pi\hbar^2}{m^2u_2^2}\,\frac{\Omega(1,2;E)}{g_2} \\
&\times \quad 4\pi\ (2\pi kT/m)^{-3/2}\ u_2^2\, \exp(-mu_2^2/2kT)\, du_2 \qquad (4.28)\\
&= \quad n_e n_2 \left(\frac{2\pi}{kT}\right)^{1/2} \frac{\hbar^2}{m^{3/2}}\frac{1}{g_2} \\
&\times \quad \left\{ \int_0^\infty \Omega(1,2;E)\exp(-E/kT)\, d\left(\frac{E}{kT}\right)\right\} \qquad (4.29)\\
&\equiv \quad n_e n_2 \left(\frac{2\pi}{kT}\right)^{1/2} \frac{\hbar^2}{m^{3/2}}\frac{1}{g_2}\Gamma(1,2)\,. \qquad (4.30)
\end{aligned}$$

The penultimate step is a change of variables from u_2 to $E = (1/2)mu_2^2$, the last line defines the dimensionless energy-weighted cross section $\Gamma(1,2)$ to be the integral in $\{\}$. The collisional excitation rate is similarly

$$\begin{aligned}
n_e n_1 \int_0^\infty u_1\sigma_{12}(u_1)\, f(u_1)\, du_1 &= n_e n_1 \left(\frac{2\pi}{kT}\right)^{1/2} \frac{\hbar^2}{m^{3/2}}\frac{1}{g_1} \\
&\times \quad \Gamma(1,2)\exp(-\Delta E_{12}/kT) \qquad (4.31)\\
&\equiv \quad n_e n_1 q_{12}\,. \qquad (4.32)
\end{aligned}$$

The last line defines the net rate coefficient q_{12}. Quantum mechanical calculations are needed to compute the dimensionless $\Gamma(1,2)$ for given atoms and ions, what we will need later is that *the rate per unit volume for collisional excitations* is

$$n_e n_1 q_{12} \propto T^{-1/2}\,\exp(-\Delta E_{12}/kT)\,. \tag{4.33}$$

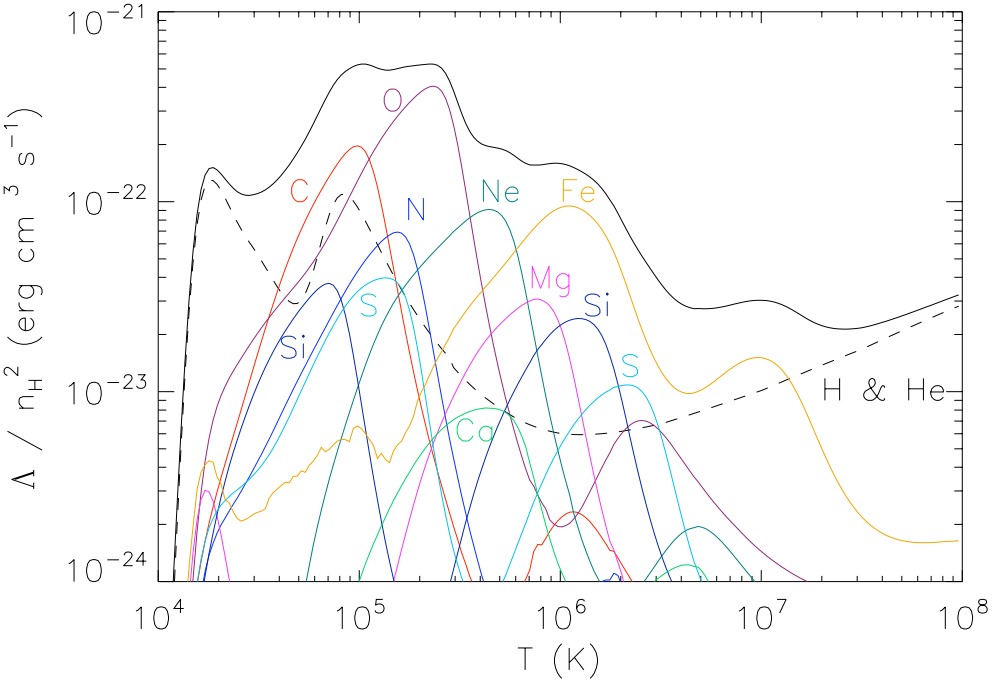

Figure 4.1: Contribution of various elements (coloured curves labeled with the corresponding element) to the total cooling rate (black curve) of a solar abundance cloud as a function of temperature T, computed using CLOUDY by Wiersma et al. (2009a). The dashed curve assumes primordial gas consisting of H and He. Collisional line cooling of electronically excited levels is not important below $T \sim 10^4 K$ as collisions are not energetic enough to excite any electronic transitions, and also not above $T \sim 10^7$ K as most atoms are completely ionised. Thermal Bremsstrahlung (from electrons) causes the upturn above $T \sim 10^7$ K.

To make sense of all this maths, consider as an example a low density gas. We assume that atoms get collisionally excited from level $1 \to 2$, but that the gas has such low density that transitions $2 \to 1$ are mostly radiative[5]. The photon with energy $h\nu_{21} = \Delta E_{12}$ may escape the system provided its optical depth is low enough. This is why forbidden lines are particularly good coolants: they have a small optical depth. A photon from an allowed transition with large optical depth may be re-absorbed in the cloud, hence fail to escape, therefore contributing little to the radiative cooling rate.

Assuming the photon escapes, the cooling rate due to such *collisionally excited lines* per unit volume is then the product of the collisional excitation rate, times the energy ΔE_{12} lost per excitation, ΔE_{12}, hence

$$\mathrm{L} = n_e \, n_1 \, q_{12} \, \Delta E_{12} \, . \tag{4.34}$$

Important coolants in the ISM are relatively abundant elements (such as C, Ni, O, Ne and Fe), see below.

Given that this cooling results from collisionally excited levels, it is easy to understand the density dependence of L, and we might have expected collisions to be ineffective when the typical collision energy is less than that required to excite the transition: when $kT \ll \Delta E_{12}$, collisional excitations are exponentially suppressed since $q_{12} \propto T^{-1/2} \exp(-\Delta E_{12}/kT)$.

For very high $T \gg \Delta E_{12}/k$, the cooling rate of our two level atom drops (slowly) $\propto T^{-1/2}$. However when applied to a real (multi-level) particle, the ion is likely to increase its ionisation state when $kT \geq \Delta E_{1\infty}$. Consider oxygen as an example. At low T, oxygen might be mostly neutral, and collisions are unimportant when $kT \ll \Delta E_{12}$, for any excited state 2 of O I. With increasing T, collisions become important, and the gas starts to cool due to excitations in O I. However for even higher T, oxygen will be mostly in the form of O II, hence cooling is now due to excitations in that ion. Similarly for even higher T, cooling is due to O III, then O IV, then O V, etc. Any particular ion will therefore be an important coolant in a given temperature range: below a minimum temperature, either most ions are in a lower ionisation state and/or the collisions cannot excite the ion, whereas at higher T

[5]Recall that the collisional transition rate $\propto n_e n_2$, whereas the radiative rate $\propto n_2$: a power of density less. Therefore collisional de-excitation will be less important than radiative ones provided the density is low enough densities.

most ions are in a higher ionisation state. Finally at sufficiently high T the species is fully ionised, and therefore no longer contributes to the cooling. This reasoning suggests it should be possible to constrain the temperature of the nebula from the relative contributions of different transitions *of the same element* to the cooling rate. We will elaborate below.

Figure 4.1 shows how different species contribute to the cooling rate as function of T. The *dominant* density dependence of the cooling rate is $\propto n^2$, as follows from Eq. (4.34), but note that the actual cooling rate does **not** have a simple power-law dependence on density. In any case this is why the dependence of cooling rate on temperature for a given density is plotted as $\Lambda \equiv (\mathrm{L}/n_{\mathrm{H}}^2)$, with the different coloured lines the contributions to Λ for a given element assuming solar abundances. Note how each species (oxygen say) adds a vaguely Gaussian shape to the overall Λ: at too low T, the collisions cannot excite the atom, at too high T the atom is completely ionised and hence does not contribute to cooling.

The full black curve is the sum of all curves, the dashed line is the case without any metals (H, He and electrons, only). At higher $T \geq 10^7$ K, the elements are almost completely ionised and the cooling rate drops. Here, an unrelated cooling mechanism, thermal Bresstrahlung due to the electrons, starts to dominate. Below $\sim 10^4$ K, line cooling due to collisional excitation of electronic energy levels, is negligible as $kT \ll \Delta E_{12}$. In practise there are other transitions not included in that figure which do allow the gas to cool[6].

The line cooling described here causes observed H II regions to be cooler than what we had naively derived in the previous section of pure H II regions.

4.2 Temperature sensitive line ratios

Kwok, §5.12
The previous section discussed how the fraction of an element that is in a given ionisation state (say O V as compared to O VI depends on temperature; however that fraction depends on density as well. Crucially however there are line ratios that *only* dependent on temperature: these are used to

[6]Fine-structure, dust, and roto-vibrational transitions in molecules for example

measure nebular temperatures, for example of H II or PNe.

An important example are the forbidden lines in [OIII], discussed in Sect. 3.2.2, which involve dipole forbidden transitions between the excited ^1S and ^1D terms to the ground state ^3P. These states can be excited through

- collisional excitation
- recombination of O IV to an excited state, followed by a radiative cascade
- absorption of a (UV) photon, followed by a radiative cascade

In most physical systems, the first process dominates. The excited state may return to the ground state either collisionally, or radiatively. We will simplify the [OIII] ion to a three level system, with $n = 1$ the ground state (^3P), and two excited levels $n = 2$ (^1D) and $n = 3$ (^1S).

Following Eq. (4.32), write the rate of collisional excitation from $n = 1 \rightarrow n = 2$ as

$$\frac{\mathrm{d}n_2}{\mathrm{d}t}|_{\mathrm{coll}} = n_1\, n_2\, q_{12}\,, \tag{4.35}$$

and similarly for the collisional excitation from $n = 1 \rightarrow n = 3$ and $n = 2 \rightarrow n = 3$. Radiative de-excitation is described in terms of the Einstein 'A' coefficients from Eq. (A.9), for example for $n = 2 \rightarrow n = 1$

$$\frac{\mathrm{d}n_1}{\mathrm{d}t}|_{\mathrm{rad}} = n_2\, A_{21}\,, \tag{4.36}$$

and similarly for $n = 3 \rightarrow n = 2$. In equilibrium, the net density of [OIII] in its three states does not change, hence all collisional excitations are balanced by the corresponding radiative and collisional de-excitations:

$$\begin{aligned} n_e n_1(q_{12} + q_{13}) &= n_2(A_{21} + n_e q_{21}) \\ &+ n_3(A_{31} + n_e q_{31}) && (4.37)\\ n_2\left[n_e(q_{21} + q_{23}) + A_{21}\right] &= n_1 n_e q_{12} + n_3\left[n_e q_{32} + A_{32}\right] && (4.38)\\ n_3\left[A_{32} + A_{31} + n_e(q_{32} + q_{31})\right] &= n_e\left[n_1 q_{13} + n_2 q_{23}\right]\,. && (4.39) \end{aligned}$$

Terms on the left represent the different channels (radiative $\propto A$, or collisional $\propto q$) by which the number density of levels 1,2,3 (top to bottom) are

depleted, whereas those on the right are the channels that produce those levels. These equations can be solved numerically and so the level population computed, once the temperature dependence of the collisional rate coefficients (the qs) and the Einstein A coefficients are known. Only two of these equations are independent, since the sum $n_1 + n_2 + n_3$ is also a constant. We can use CLOUDY to compute (or tabulate) the occupation distribution, but it is instructive to make some more approximations and obtain a (nearly) analytic relation as follows.

For the typically relatively low densities of nebulae in the ISM, these equations can be simplified by neglecting collisional de-excitations as compared to radiative ones. Note that a collisional de-excitation (for example $2 \rightarrow 1$ is $\propto n_e n_2 q_{21}$ and hence will be slower at lower density than the radiative one which is $\propto n_2 A_{21}$, one power of density less. Furthermore we can usually assume most atoms to be in the lower energy states, so that $n_3 \ll n_2 \ll n_1$.

Under these *approximations* the last two equilibrium relations become

$$n_2 A_{21} \approx n_1 n_e q_{12} \tag{4.40}$$

$$n_3 \left[A_{32} + A_{31}\right] \approx n_1 n_e q_{13}\,, \tag{4.41}$$

therefore

$$\frac{n_2}{n_3} \approx \frac{q_{12}}{q_{13}} \frac{A_{32} + A_{31}}{A_{21}}\,. \tag{4.42}$$

Applying this to [OIII], Fig. 3.2 shows the $n = 3 \rightarrow 2$ line has $\lambda = 436.3$ nm, whereas the $n = 2 \rightarrow 1$ is a doublet consisting of $\lambda = 500.7$ and 495.9 nm. The intensity of a given line say $n = 3 \rightarrow 2$ is $I(3 \rightarrow 2) \propto n_3 A_{32} h\nu_{32}$: the product of the rate at which the transition occurs, times the energy of the photon. Therefore the ratio of emission line strengths

$$\frac{I(436.3)}{I(500.7 + 495.9)} = \frac{n_3 A_{32} h\nu_{32}}{n_2 A_{21} h\nu_{21}} \tag{4.43}$$

$$= \frac{q_{13}}{q_{12}} \frac{A_{32}}{A_{32} + A_{31}} \frac{h\nu_{32}}{h\nu_{21}}\,. \tag{4.44}$$

Given the temperature dependence of the qs (Eq.4.33), the implies this line ratio has a temperature dependence $\propto \exp(-\Delta E_{13}/kT)/\exp(-\Delta E_{12}/kT) \propto$

$\exp(-\Delta E_{23}/kT)$. Measuring the line intensity ratio and using the atomic values for O III yields the temperature of the nebula. This method is an important temperature diagnostic at the low densities typically encountered in the ISM. Hopefully this is obvious but nevertheless it is import to note that intensity ratio is independent of the oxygen *abundance*.

4.3 Exercises

1. Demonstrate that Eq. (4.2) indeed solves Eq. (4.1).

2. Solve equation (4.16). Assume $\lambda = 0$ corresponds to $x = 1/2$, the location in the nebula where half the gas is ionised. Compute the level of ionisation at the location where $\lambda = 10$. Verify that the gas is indeed mostly neutral there, and hence that the transition from mostly ionised to mostly neutral occurs over a distance small compared to the Strömgren radius. [Hint: see Dyson & Williams p. 71]

Chapter 5

Line formation

Dyson & Williams, 5.1–5.3, Kwok §5.12

- Thomson scattering
- Natural line profile
- Thermal broadening
- Voigt profile

Previously we described how the cross-section for radiative bound-bound and bound-free transitions in atoms can be computed. Here we discuss in more detail the shape of the resulting absorption/emission lines.

In the previous chapter we described how an atom/ion can be put in an electronically excited state by a collision, for example a hydrogen's electron can end-up in the $n = 2$ state, rather than the $n = 1$ ground state. When the electron transitions radiatively to $n = 1$, it emits a Lyman-α photon, which we may be able to detect. For example an H II region will emit a whole series of hydrogen lines, including Lyman-α, Lyman-β, but also Balmer (H)-α, H-β etc. The relative strength of the observed lines does not just depend on atomic physics, but also on the physical properties of the emitting system (e.g. the temperature, density and extent of the H II region) which determines the optical depth of the emitted photon. With most hydrogen atoms in general in the ground state, photons from the Lyman-series encounter a (much) higher optical depth than those of say the Balmer series, hence H II regions are particularly bright in H-α. We also remind the reader that for the same reason forbidden lines can be very strong, e.g. a

very large fraction of the luminosity of PNe comes out in the forbidden OIII line discussed previously: the low density implies that radiative transitions dominate, the small optical depth allows the emitted photon to escape the nebula.

Intervening gas can also absorb light from a background source. This allows us to study the Milky Way's ISM in the spectra of background stars, say, and the gas surrounding galaxies (the circum-galactic medium, CGM) and in between the galaxies (the intergalactic medium, IGM) in the spectra of background quasars. Intervening hydrogen absorbs in the Lyman-series (and *much* weaker in the Balmer series). In this case the atom is radiatively excited. When it de-excites, the resulting photon is in general emitted in a different direction. This 'absorption process' is therefore more accurately described as *scattering*, though most of the literature refers to this as absorption. A hopefully familiar example showing that it is better to think of this as scattering rather than absorption is the scattering of Sun light off molecules (mostly N_2 and O_2) making the sky appear blue - and dimming the Sun.

To describe the relevant physics we start with a simple case.

5.1 Thomson scattering

Scattering off electrons is called Thomson scattering[1]. The electric field $\mathbf{E} = \mathbf{E_0}\sin(\omega_0 t)$ of the incoming electro-magnetic radiation will exert an oscillating force on the charged particle of charge e and mass m. This will make the particle oscillate and radiate energy. In the classical limit, the fraction of light scattered is independent of frequency, and the cross section is the Thomson cross section:

$$\sigma_T = \frac{8\pi}{3}\,(\frac{e^2}{mc^2})^2\,, \tag{5.1}$$

which a value of $\sigma_T \approx 6.625 \times 10^{-25}\,\mathrm{cm}^2$ for the electron. Note that the interaction strength $\propto 1/m^2$ so Thomson scattering on protons is much less efficient. Comparing with Eq. (1.19) yields the optical depth for Thomson scattering, for a given path length ds through gas with electron density n_e,

$$d\tau = \sigma_T\, n_e\, ds\,. \tag{5.2}$$

[1] Rybicki & Lightman, section 3.4

5.2 Scattering by a harmonically bound particle

The[2] next simplest system is scattering by a charged particle bound in a harmonic potential well. The equation of motion of such a system is

$$m\ddot{x} = -m\omega_0^2 x + m\tau\,\dddot{x} + eE_0 \sin(\omega t)\,. \tag{5.3}$$

Here, ω_0 is the frequency of the harmonic oscillator, τ (not to be confused with the optical depth) represents a damping term, and the external radiation field is driven with frequency ω. The cross section for scattering in this case is frequency dependent: the particle will radiate much more efficiently when in resonance with the incoming radiation, *i.e.* when $\omega \approx \omega_0$. The cross section now has the shape of a line with a given width and a shape given by

$$\begin{aligned}
\sigma(\nu) &= \frac{\pi e^2}{mc}\,\frac{\gamma/2\pi}{(\nu-\nu_0)^2+(\gamma/2)^2} \\
&= c\sqrt{\frac{3\pi\sigma_T}{8}}\,\frac{\gamma/2\pi}{(\nu-\nu_0)^2+(\gamma/2)^2} \\
&= \sigma_0\,\phi(\nu) \\
\sigma_0 &\equiv c\sqrt{\frac{3\pi\sigma_T}{8}} = 0.0265\,\mathrm{cm^2\,s^{-1}}\,,
\end{aligned} \tag{5.4}$$

where $\gamma = \omega_0^2\tau$. Note that $\int_0^\infty \phi(\nu)d\nu = 1$, so that σ_0 is the net cross section of the transition, and ϕ describes the line *shape*,

$$\phi(\nu) = \frac{\gamma/2\pi}{(\nu-\nu_0)^2+(\gamma/2)^2}\,. \tag{5.5}$$

This functional form is called a **Lorentz profile**.

5.3 Scattering by bound transitions

We described in Chapter 3 how the interaction of photons with bound states of an atom[3], for example the electronic states of the Hydrogen atom, can

[2]Rybicki & Lightman, section 3.6

[3]Rybicki & Lightman section 10

be computed by treating the atom quantum mechanically, but the radiation field classically. Let's take hydrogen as an example.

The electronic energy levels have energies $E_n = -E_0/n^2$, where $E_0 = 13.6$ eV is the binding energy. The energy of the photon emitted by a bound-bound transition between levels n_2 and $n_1 < n_2$ is $E_{12} = E_0(1/n_1^2 - 1/n_2^2)$. Alternatively, an atom in energy state n_1 can absorb a photon of energy E_{12} to get into the excited state n_2. Useful to remember are also the corresponding wavelengths and f-values

Table 5.1: f-values and wavelengths for the first three Lyman-series lines in the Hydrogen atom, and for the Lyman-limit

line name	λ (Å)	f-value
Lyman-α ($n = 1 \rightarrow 2$)	1215.6701	0.4164
Lyman-β ($n = 1 \rightarrow 3$)	1025.7223	0.0791
Lyman-γ ($n = 1 \rightarrow 4$)	972.7431	0.02899
Lyman-limit ($n = 1 \rightarrow \infty$)	911.8	

The cross section of this process is frequency dependent and strongly peaked at the resonance frequency $\nu_0 = \nu_{12} = E_{12}/h$,

$$\sigma(\nu) = \frac{\pi e^2}{mc} f \frac{\gamma/2\pi}{(\nu - \nu_0)^2 + (\gamma/2)^2} . \tag{5.6}$$

The expression is very similar to that of the damped, driven harmonic oscillator, Eq. (5.2). The cross section is reduced by a factor f, imaginatively called the f-value of the transition, and given in the table for the first Lyman transitions. The natural line width γ of the transition results from the Heisenberg uncertainty principle, and is related to the Einstein coefficients, see e.g. Hilborn (2002).

The f-value for a given transition is an integral over the product of the wave-functions of the n_1 and n_2 states, and values for the first three Lyman transitions, (1–2, 1–3, 1–4) are given in Table 5.1. They can be computed exactly for the Hydrogen atom.

Because the shape of the line is so similar to that of the driven, damped oscillator, the extended line wings of the Lorentz profile are often called **damping wings** (since for the harmonic oscillator they result from the damping

term), and the cosmological structures in which you can detect this wing, are called **damped Lyman-alpha systems**.

5.4 Line-broadening

5.4.1 Thermal broadening

According to the Maxwell-Boltzmann distribution, the fraction of particles of mass m, in a gas in local thermodynamic equilibrium, with velocity in the $z-$direction in the interval $[v_z, v_z + dv_z]$, is

$$N(v_z) = \frac{1}{\sqrt{2\pi kT/m}} \exp(-v_z^2/(2kT/m)) \,. \tag{5.7}$$

The scattering cross section of each of these particles peaks at the Doppler shifted frequency $\nu = \nu_0 \,(1 + v_z/c)$. Therefore the gas absorbs not just at the resonant frequency ν_0, but the Doppler motion of the atoms broadens the line. The line-profile can be found by realizing that the fraction of photons absorbed at the Doppler shifted frequency ν is just proportional to the fraction of particles with velocity v_z: $\phi(\nu)\, d\nu = N(v_z)\, dv_z$, hence the line-shape

$$\begin{aligned} \phi(\nu) &= N(v_z) \frac{dv_z}{d\nu} \\ &= \frac{c}{\nu_0} \frac{1}{\sqrt{2\pi kT/m}} \exp(-(\nu - \nu_0)^2/(2\nu_0^2 kT/mc^2) \,, \end{aligned} \tag{5.8}$$

which can be rewritten as

$$\begin{aligned} \phi(\nu) &= \frac{1}{\sqrt{\pi\, \Delta\nu_D^2}} \exp(-(\nu - \nu_0)^2/\Delta\nu_D^2) \\ \Delta\nu_D &= \frac{\nu_0}{c} \sqrt{\frac{2kT}{m}} \,. \end{aligned} \tag{5.9}$$

5.4.2 Voigt profile

Heizenberg's uncertainty principle caused broadening of spectral lines according to the Lorenz profile, Eq. (5.6). In a gas with thermal motions, each atom

moving with velocity v_z will produce an absorption line with a Lorentzian shape. The net profile is then the *convolution* of these two profiles:

$$\phi(\nu) = \int_{-\infty}^{\infty} dv_z \frac{1}{\sqrt{2\pi kT/m}} \exp(-v_z^2/(2kT/m)) \frac{\gamma/4\pi^2}{(\nu - \nu_0(1+v_z/c))^2 + (\gamma/4\pi)^2} \,. \tag{5.10}$$

Define

$$\begin{aligned} u &\equiv \frac{\nu - \nu_0}{\Delta\nu_D} \\ a &\equiv \frac{\gamma}{4\pi\Delta\nu_D} = \frac{\Delta v_{\rm H}}{b} \,, \end{aligned} \tag{5.11}$$

where $\Delta v_{\rm H} = 6.06076 \times 10^{-3} {\rm km\,s^{-1}}$ is the natural line width, and $b = \sqrt{2kT/m} = c\Delta\nu_D/\nu_0$ is the Doppler width.

The line profile is then

$$\begin{aligned} \phi(\nu) &= \frac{1}{\Delta\nu_D} H(a,u) \\ H(a,u) &= \frac{a}{\pi} \int_{-\infty}^{\infty} \frac{\exp(-y^2)}{a^2 + (u-y)^2} \, dy \,. \end{aligned} \tag{5.12}$$

$H(a,u)$ is called the *Voigt* function, and the line-shape $\phi(\nu)$ is called the Voigt profile.

The width γ of the natural line profile is much narrow than the Doppler width in most cases, and the net line profile is then indistinguishable from a Gaussian. However note that the Voigt profile is dominated by the Lorenzian in the wings. Therefore the very strong lines in a DLA, for which the central optical depth is very high, still have significant absorption far from the line centre, and the line profile there is dominated by the Lorentzian shape.

As an example, consider a slab of gas, with HI column density $N_{\rm HI}$, moving with respect to the observer with velocity v_0. This will produce an absorption line centered at wave length $\lambda = \lambda_0(1 + v_0/c)$, and the amount of absorption at velocity v is

$$\tau(v) = c\sqrt{\frac{3\pi\sigma_T}{8}} \, f \, \lambda_0 \, \frac{N_{\rm HI}}{\Delta\nu_D} \, H(a, u + \frac{\nu_0 v}{c\Delta\nu_D}) \,. \tag{5.13}$$

The factor $c\sqrt{3\pi\sigma_T/8}\, f = \pi e^2 f/mc$ is the net cross section (cfr Eq. 5.6), it is multiplied by the column density $N_{\rm HI}$, and the Voigt profile is shifted from u to $u + (\nu_0 v/c\Delta\nu_D)$ to take into account the cloud's Doppler shift.

Chapter 6

The intergalactic medium

Consider a bright and distant source such as a quasar or the optical radiation associated with a γ-ray burst. The light of such a source travels through most of the Universe to get to us - will it be absorbed by the intergalactic medium (IGM) - the gas in between the galaxies?

6.1 Gunn-Peterson trough

Suppose the IGM were homogeneous, with proper density $n(z)$ (where z is the redshift). What fraction of the light from a distant quasar would be scattered by the hydrogen atoms along the line of sight?

Let $\rho_c = 3H_0^2/(8\pi \mathrm{G})$ be the critical density, and $y \approx 0.24$ be the Helium mass fraction, then the mean number density of hydrogen atoms at redshift $z = 1/a - 1$ is

$$\begin{aligned} n_H(a) &= n_0\, a^{-3} \\ n_0 &= 1.7 \times 10^{-7}\, \frac{\Omega_b h^2}{0.02}\, \mathrm{cm}^{-3}\,. \end{aligned} \tag{6.1}$$

Observed and emitted frequencies are related by $\nu_{\rm obs} = \nu_{\rm em}/a$, where $a = 1/(1+z)$ is the expansion factor. For a photon in an FRW metric, the proper lenght $dl = cdt = cda/\dot{a} = cda/aH(a)$, and at redshifts ≥ 1 we can use the Einstein-de Sitter approximation for the Hubble constant, $H(a) \approx H_0 \sqrt{\Omega_m}\, a^{-3/2}$, where H_0 is Hubble's constant today, and $\Omega_m \approx 0.3$

the matter density. We will also assume that the Lyman-alpha cross section is very peaked, so that $\sigma(\nu) \approx \sigma_0\,\delta(\nu-\nu_0)$, where δ is Dirac's delta function.

Recall that by definition

$$\int_a^b f(x)\,\delta(x-x_0)\,dx = f(x_0)\,, \tag{6.2}$$

when x_0 is in the interval $[a,b]$, and zero if x_0 is outside that interval.

The optical depth at frequency ν, when observing a source at emission redshift $z_{\rm em} = 1/a_{\rm em} - 1$, is then

$$\begin{aligned}
\tau(\nu) &= \int_{a_{\rm em}}^1 \sigma(\nu/a)\,n(a)\,\frac{c\,da}{a\,H(a)} \\
&= \frac{\sigma_0\,n_0\,c}{H_0\,\sqrt{\Omega_m}}\int_{a_{\rm em}}^1 \delta(\nu/a-\nu_0)\,a^{-5/2}\,da \\
&= \frac{\sigma_0\,n_0\,c}{H_0\,\sqrt{\Omega_m}}\,\frac{a^{-3/2}}{\nu_0} \\
&\approx 13000\,h^{-1}\,\frac{\Omega_b\,h^2}{0.02}\,(1+z)^{3/2}\,.
\end{aligned} \tag{6.3}$$

(Note that $\tau(\nu) = 0$ for $\nu \le \nu_0/a$.) Therefore the optical depth at redshift $z \sim 3$ is about $\tau(z=3) \approx 1.5\times10^5$. Clearly the reason we see any flux below the Lyman-alpha emission line is that the Universe is very highly ionised, so that $n_{\rm HI}/n_{\rm H} \approx 1\times10^{-5}$. An example of a $z\sim3$ quasar spectrum is shown in Fig.6.1.

The photo-ionisation of the IGM also heats the gas, as discussed in the next section.

6.2 Thermal history of the IGM

The IGM is photo-ionised and heated by radiation from hot stars and quasars. Before any of the these first sources had formed, the IGM was neutral following recombination. The epoch when the first sources formed and started ionising the gas around them until eventually the IGM was fully ionised is referred to as the epoch of reionisation.

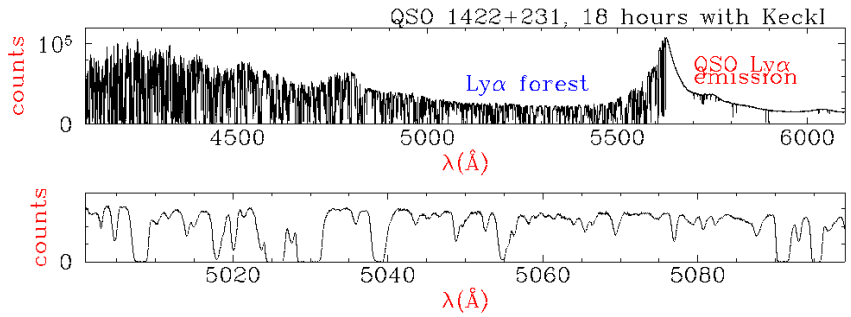

Figure 6.1: Observed spectrum of QSO 1422, one of the brightest quasars at $z \approx 3$. The broad Lyα emission line of the quasar is indicated in the top panel, as is the dense set of intervening absorption lines that correspond to the Lyα forest. The bottom panel is a zoom.

6.2.1 Before reionisation

At recombination, the free electron keep the gas temperature tightly coupled to the CMB temperature, $T_{\rm gas} \approx T_{\rm CMB}$. The expanding Universe cools the gas adiabatically. For an adiabatic gas, $p \propto \rho^\gamma$ hence $T \propto \rho^{\gamma-1}$. Since $\rho \propto (1+z)^3$, this gives

$$T \propto (1+z)^{3(\gamma-1)}\,, \tag{6.4}$$

hence $T \propto (1+z)$ for the CMB ($\gamma = 4/3$) and $T \propto (1+z)^2$ for the Hydrogen/Helium IGM ($\gamma = 5/3$). Close to recombination, residual electrons left over from recombination can keep $T_{\rm gas} \approx T_{\rm CMB}$ but below $z \approx 100$ the gas cools faster than the CMB.

6.2.2 After reionization

We have seen that the IGM is very highly ionized at $z \sim 3$. Suppose therefore that the IGM is bathed in the ionizing flux of a population of sources. The argument that lead to Olbers' paradox suggests that the typical neutral hydrogen photon sees very many sources, therefore we will assume that the flux

of this putative UV-background is (nearly) uniform, but redshift dependent. Let $J(\nu, z)$ be this UV-flux.

The photo-ionization rate that results from this flux is

$$\Gamma(z) = \int_{\nu_{\rm th}}^{\infty} \frac{4\pi J(\nu, z)}{h\nu}\, \sigma(\nu)\, d\nu\,, \tag{6.5}$$

The first factor converts energy flux to photon flux, the second factor is the photo-ionization cross section for Hydrogen, and we only integrate over those frequencies for which $h\nu > E_0$, the binding energy for the Hydrogen atom. The rate of change of the neutral density is then (assuming pure hydrogen)

$$\frac{dn_{\rm HI}}{dt} = \alpha n_e\, n_{\rm HII} - \Gamma_e n_e\, n_{\rm HI} - \Gamma\, n_{\rm HI}\,, \tag{6.6}$$

The firsts term is the recombination rate (where the atomic constant α is a function of temperature), the second term is the collisional ionization rate (where an electron collides with a neutral hydrogen atom), and the last term is the photo-ionization rate. Neglecting collisions, and defining the neutral fraction $x \equiv n_{\rm HI}/n_H$, we find that the equilibrium neutral abundance

$$x = \frac{\alpha n_H}{\Gamma} \tag{6.7}$$

when $x \ll 1$ (which is true in the IGM below $z = 6$).

A photon needs more energy than $E_0 = 13.6$ eV to ionized Hydrogen. The excess energy of the photon goes into the kinetic energy of the electron, and this excess energy represent a heating term. The heating rate per ionization is

$$\epsilon(z) = \int_{\nu_{\rm th}}^{\infty} \frac{4\pi J(\nu, z)}{h\nu}\, \sigma(\nu)\, (h\nu - h\nu_{\rm th}) d\nu\,, \tag{6.8}$$

since the excess energy per photo-ionization is $(h\nu - h\nu_{\rm th})$.

The heating rate is therefore

$$\rho \frac{du}{dt} = \epsilon\, n_{HI}\,. \tag{6.9}$$

This shows that the photo-heating rate is

$$\frac{dT}{dt} \propto \frac{\alpha\epsilon}{\Gamma}\, n_H\,. \tag{6.10}$$

Note that this is independent of the amplitude of the UV-background J, since both ϵ and Γ are proportional to J. The heating rate is proportional to density.

The photo-ionization cross section, $\sigma(\nu)$, is sharply peaked around $\nu = \nu_{\rm th}$ falling as ν^{-3} at higher frequencies. This means that most photo-ionizations are by low energy photons, and hence the photo-heating rate is low. However, when the gas is ionized from neutral, then all photons ionize a Hydrogen atom, and the heating rate is now

$$\epsilon(z) = \int_{\nu_{\rm th}}^{\infty} \frac{4\pi\, J(\nu, z)}{h\nu}\, (h\nu - h\nu_{\rm th}) d\nu\,, \tag{6.11}$$

The consequence is that the gas gets heated very strongly during reionization, with a post-reionization temperature that is only weakly dependent on density. After reionization, the gas cools almost adiabatically, with a small heating from the UV-background.

Gas at lower densities cools more, and gets heated less. This introduces a temperature-density relation that looks like $T \propto \rho^{\gamma-1}$, where $\gamma \approx 1$ immediately after reionization, and γ increasing at lower redshifts.

Chapter 7

Abundance evolution

- Stellar evolution of low and intermediate-mass stars, massive stars, and binary stars, their element production and life-times
- s and r-process, supernova types
- Stellar initial mass function
- Evolution of abundances in a *closed box* system
- G-dwarf problem and its resolution

7.1 The production of the elements

Burbridge, Burbridge, Fowler & Hoyle (*Synthesis of the elements in stars*, Burbridge et al., 1957) and independently Cameron (*On the origin of the heavy elements*, Cameron 1957), published the first papers on the origin of the elements, and Fowler received the Nobel prize for this in 1983, together with Chandrasekhar. To what extent can what we know about stars and their evolution explain the abundance of elements[1] is the MW stars, or in particular in the Sun as seen for example in Fig. 7.1?

[1] Commonly, yet erroneously, called *chemical* evolution.

7.1.1 Basics of stellar structure & evolution

The structure of a star is governed by the equations of hydrostatic equilibrium and the equation of state of the stellar material, and makes in fact no reference to energy production[2]. All three of density, temperature and pressure increase towards the centre of the star. Given the Coloumb barrier, this means that hydrogen fusion can only operate via quantum tunnelling in the central 'core'. Stars burning hydrogen there are *Main Sequence* stars, and the main sequence life-time varies $\propto M/L \approx M^{-2}$ given the strong dependence $L \propto M^{\alpha}$ with $\alpha \approx 3$ between mass and luminosity. Massive stars hence have short main sequence life-times.

When the fuel in the core is used-up, the star will contract and hence heat-up. Sufficiently massive stars can start burning helium to carbon, and also hydrogen to helium in a shell: these are Red Giants. Very massive stars can keep going along this path of burning previous ashes at higher temperatures[3] following contraction, and end-up with a layers of more massive elements towards their centres: this evolution stops once the cores gets converted to Fe, for which the nucleons are most tightly bound (Fig. 7.2): fusion can then no longer compensate for the energy the star loses through radiation. The layered structure of such a star is depicted in Fig.7.3

7.1.2 AGB stars

Intermediate-mass stars with $M \leq 10 M_{\odot}$ build-up a C/O core following He burning, surrounded by a He-burning shell, itself embedded in a H-burning shell. In this *Asymptotic Giant Branch* stadium, the envelope is convective, and pulsations driven by the He and H burning shells can convectively bring core material to the surface in a process called *dredge-up* . Grains forming in the cool envelope drive a strong wind, enriched with burning produce.

The burning shells are strong neutron sources, produced by C or Ne burning. Neutrons are not hampered by a Coulomb barrier and their interaction

[2]We will not discuss degenerate stars here.

[3]A little thought suggests that such a nuclear burning sequence builds-up 'α elements'that differ by one He (α particle), so C, O, Ne, Mg, Si, S, Ar, Ca, ..., etc up to Fe

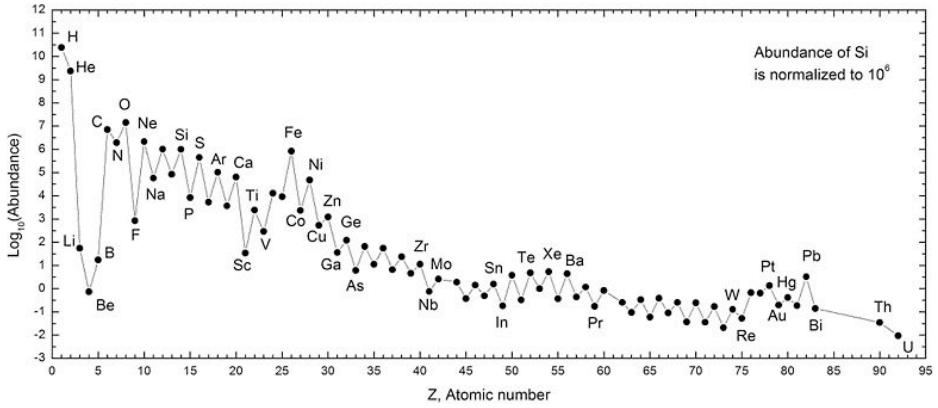

Figure 7.1: Abundances of elements for the Sun, normalised to that of Si. Note the striking 'odd-even' pattern,

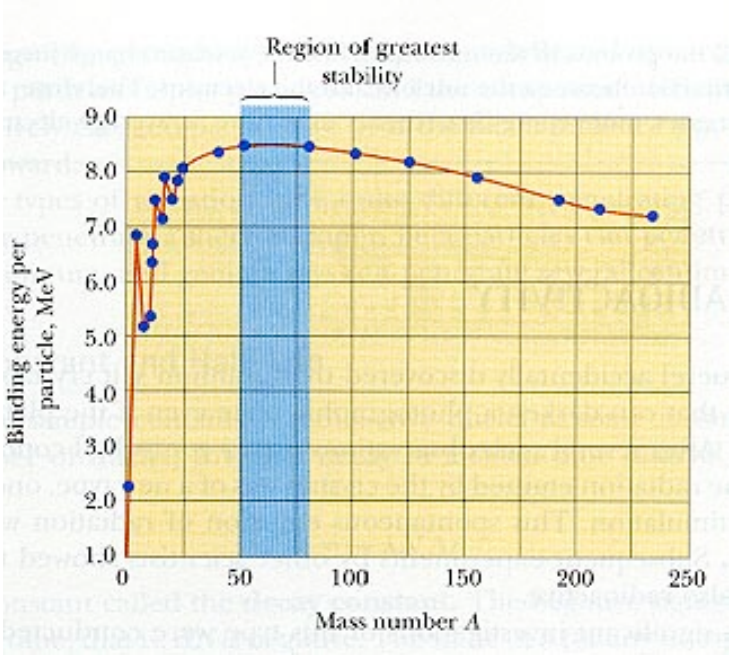

Figure 7.2: Binding energy per nucleon, from Alward.

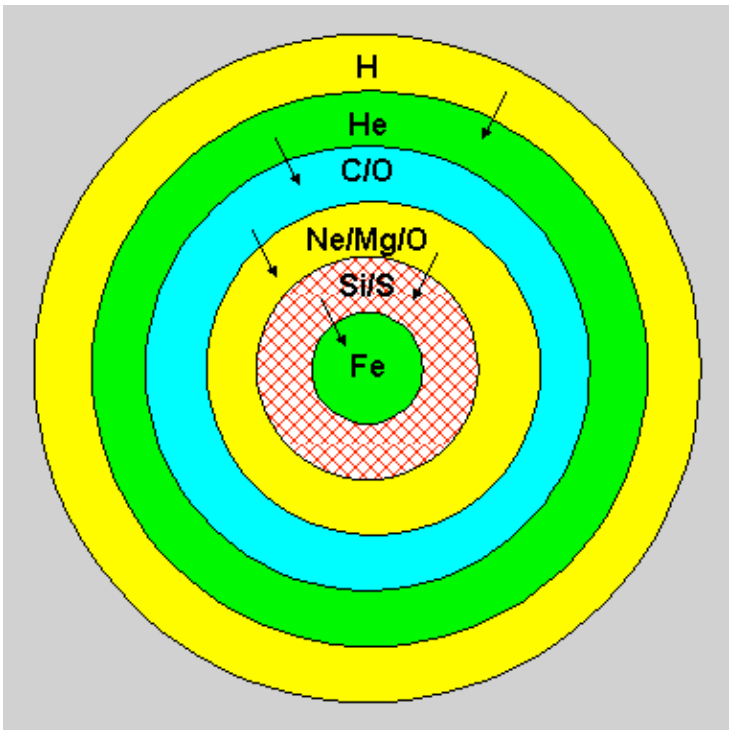

Figure 7.3: Structure of type II SNe progenitor, from www.hacastronomy.com

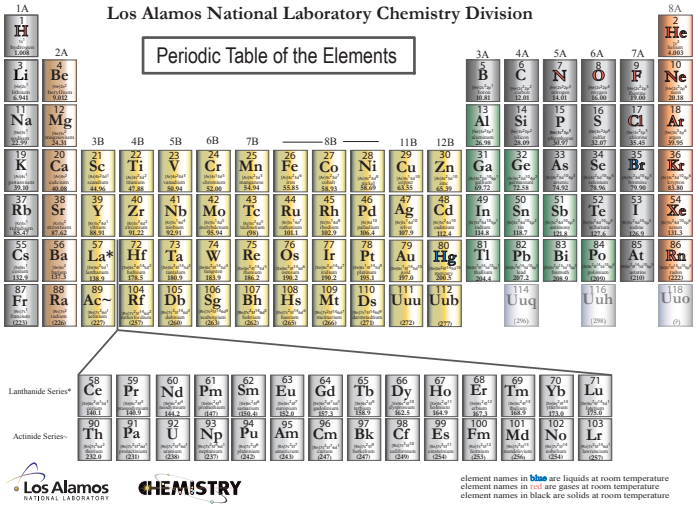

Figure 7.4: Periodic table of the elements, from Los Alamos.

with nuclei builds various isotopes. An n-rich nucleus can change type by β-decay if left enough time before being bombarded by neutrons again. Given that the n-source is weak compared to that in the massive stars discussed later, this process is called the *s-process* (for *s*low), and the abundance patterns of stars such as the Sun as shown in Fig. 7.1 was used to postulate its existence.

Most of the enriched mantle of AGB stars is flung into space reaching its planetary nebula stage, and AGB stars are thought to produce a large fraction of the low-mass elements, in particular about 1/2 of all C is thought to be produced by AGB stars.

Very-low mass stars have such long life times that they do not enrich the ISM at all: they are effectively sinks of mass and metals.

7.1.3 Type II SNe

Type II supernovae are characterised by H-lines in their spectra, and they are thought to be the end state of the evolution of massive stars, $M \geq 10 M_{\odot}$, although stars of slightly lower mass may explode as well. Pre-explosion HST images of sites where later a type II SN was seen show indeed the presence of a very luminous blue and hence massive star.

These massive stars have shells of increasingly massive fuel ashes at increasing depth, Fig.7.3. The burning produce typically differ by one α particle (or He-nucleus), and hence they are very *alpha-rich*; see Fig. 7.4. The star is also a very strong neutron source, and this r (for *r*apid) process also produces elements through β-decay as well as isotopes. Once the nucleus has finished fusing to Fe the star has no more exothermic fusion options and the core collapses, cooling by emitting neutrino's.

Exactly how this starts a thermo-nuclear deflagration, i.e. an explosion where the explosion itself produces more energy as it travels through the star by fusion, is unclear, and numerical simulations model this by compressing the material by hand. In fact, uncertainty of where the explosion actually starts is one of the main uncertainties in computing the yields (*i.e.* how much elements of each type are produced), together with the problem of accurately resolving the burning front. The explosion releases a large fraction

of the initially α-enriched material in its surroundings, together with elements produced by fusion. In particular such explosions are thought to have produced the more exotic elements such as for example uranium. Although the star has lots of Fe in its core, only a small fraction of that Fe makes it into the surroundings, and most gets crushed into the SN remnant, either a neutron star or black hole.

Most of the energy ($\sim$ 99 per cent) of the explosion is in the form of neutrinos, but the remaining 1 per cent is still $\sim 10^{51}$ ergs, and a vast source of energy for the stirring the ISM, and potentially even removing a significant amount of baryons out of the galactic potential well in the form of a *galactic wind*.

In short, type II SNe are the end stages of massive stars, and are the main source of α-elements, such as O, Mg, Ne and Si, and possible all of the more exotic elements such as U.

7.1.4 Type I SNe

Type I SNe have spectra devoid of H-lines, but it is not so clear what are their progenitors. One theory is that Type Ia SNe result from binary evolution: if one component of a binary star has already reached its C/O WD stage, then mass transfer from its companion, either a MS or itself a WD, may push the star above the Chandrasekhar mass limit. This makes the star unstable and initiates the explosion. This model would explain why all type Ia are so similar, since the progenitor is basically always of the same mass and nearly the same composition. Chevalier (Nature 260, 689, 1976) suggested this kind of explosion as the main source of Fe in the Universe.

7.1.5 Summary

There are three main channels for element production: AGB stars, and type I and type II SNe. These channels produce elements in very different ratios, and, given the very different life-times for AGB and type I SNe on the one hand (low and intermediate mass stars) as compared to type II SNe (massive stars), at very different times *after the formation of these stars.*

7.2 Ingredients for computing the evolution of abundances

Since stars of different masses produce elements with a very different abundance pattern (for example mostly C for AGB star, α rich for massive stars, Fe-rich for type I SNe) we need to know how many stars of given mass a given population has formed to predict the abundance pattern of its ejecta - and hence of newly formed stars. Also, the life-times and hence the rate at which these stars produced these elements, are very different: both AGB and type I SNe have low/intermediate-mass star progenitors, and these evolve very slowly as compared to the massive star progenitors of type II SNe. This then determines the required ingredients for computing the abundance evolution of a galaxy.

Stellar Initial Mass Function
The initial mass function (IMF) $\Phi(M)$ is defined such that $M\Phi(M)\,dM$ is the fraction of mass in stars of mass $[M, M + dM]$. This function is difficult to compute theoretically, with modern theories suggesting that its low-mass end may be set by the properties of the turbulence observed in the molecular clouds where stars form. Observationally Φ can only be measured in nearby star forming regions, since the low-mass tail consists of low-mass and hence faint stars.

There is no a priori reason why the lowest-mass objects to form in a GMC should have masses high-enough to be able to ignite H-fusion in their centres[4] (thought to be $\approx 0.1\,M_\odot$), but an interesting suggestion is that feedback from fusion once the stars switches on ends mass accretion - possibly explaining why a collapsing GMC makes stars rather than inert low-mass clumps. Similarly there seems no obvious reason why objects should stop accreting once they reach masses $\sim 100\,M_\odot$ above which stars are thought to become unstable. In any case, no observed stars have masses significantly higher than $100M_\odot$. The IMF is thus to be measured between $M_{\rm min} \approx 0.1M_\odot$ and $M_{\rm max} \approx 100M_\odot$. Two fits to observations are often used, the *Salpeter*

[4]In fact some GMCs seem to contain freely floating *planets* with mass below the minimum mass for H-fusion.

IMF, which has $M\Phi(M) \propto M^{-1.35}$, and the *Chabrier* IMF, which has

$$M\Phi(M) \propto \exp\left[-(\log(M/M_c))^2/2\sigma^2\right] \quad \text{for } M \leq 1M_\odot \tag{7.1}$$

$$\propto M^{-1.3} \quad \text{for } M \geq 1M_\odot\,. \tag{7.2}$$

where $M_c = 0.079M_\odot$ and $\sigma = 0.68$. The Salpeter IMF is mostly misquoted in the literature, stemming from the fact that Salpeter only measured stars above $1M_\odot$ (because lower mass stars were too faint to be observed in those days), and hence his proposed power-law only applies above that limit. Chabrier used more sensitive observations and extended Salpeter's measurements to lower masses. But in the literature, what is generally (but wrongly) meant by a Salpeter IMF is a power-law mass function up to the lowest mass, in contrast to a Chabrier IMF which has a cut-off at low M.

Stellar life-times
The life-expectancy of (single) stars of given mass and abundance can be computed from stellar evolution calculations. Main uncertainties are due to mixing (where fuel from the surroundings is mixed into the burning layer providing more fuel and hence prolonging the life-time), the presence of rotation and of magnetic fields. The models can also predict how much elements of each type are released into the surroundings at the end of the stellar life, either by winds and PNe, or the ensuing SNe explosion.

The star formation history
Suppose from an initial mass M of gas, a fraction y is converted into stars with known IMF. Given the stellar life-times, and given the amount of elements those stars release as a function of time, it should be possible to compute how the abundance pattern of the remaining gas changes in time. If stars were to form at later times from this enriched gas, they would have a higher abundance (metallicity) than the original stars. We would need to know the whole star formation history (*i.e.* how many stars formed at each time, and of each mass) to compute the abundance pattern at later times. Numerical simulations of the formation of galaxies are now able to combine star formation with stellar evolution to make such predictions. However the uncertainties in some of the basic ingredients such as yields are still sufficiently large that the predictive powers of such calculations are still small. See Wiersma et al. (2009b) for a detailed description of the yields and life-times used in EAGLE. Here we give a rather simplistic example.

7.3 Stellar abundance evolution in a closed box model

A simple model where we can compute the evolution of the abundance pattern is a *closed box* model, of which the total mass M does not change as some fraction of its gas gets converted into stars at each time. Let M be that initial (and constant) mass. At time t, an amount $dM'_\star$ of stars forms in a time interval dt, of which $dM_\star$ goes into long-lived stars (or remnants), and $dM''_\star$ goes into short lived stars. We make the approximation that these stars die instantaneously, and return their mass, enriched by burning, to the gas.

Let M_Z be the total mass in metals in the gas phase, and define its metallicity $Z \equiv M_Z/M_g$, with M_g the current gas mass. As the newly formed stars die, they enrich their surroundings with metals by an amount

$$dM_Z|_{\rm gain} = p\, dM_\star\,, \tag{7.3}$$

where p is called the *yield*. The long-lived stars are sinks of metals, and as they form they lock metals away for ever. This decreases the metal mass in gas by

$$dM_Z|_{\rm loss} = -Z\, dM_\star\,, \tag{7.4}$$

where we assume they form with metallicity Z. How does Z evolve? From the definition of $Z = M_Z/M_g$, we get

$$dZ = \frac{dM_Z}{M_g} - Z\,\frac{dM_g}{M_g} = -p\,\frac{dM_g}{M_g}\,, \tag{7.5}$$

since $dM = 0 = dM_\star + dM_g$ in our close box model. Therefore

$$Z = Z_i - p\log(M_g/M)\,, \tag{7.6}$$

where Z_i is the initial abundance.

Stellar abundances
For $Z_i = 0$, the amount of stars with metallicity below some value Z_1 is

$$M_\star(<Z_1) = M - M_g(<Z_1) = M(1-\exp\left(\frac{-Z_1}{p}\right))\,. \tag{7.7}$$

This is the distribution of stars as function of their metallicity, and this simple model works surprisingly well for bulge stars.

The derivative of the previous equation yields

$$dM_\star(Z) = \frac{M}{p} \exp(-Z/p)\, dZ\,, \tag{7.8}$$

and therefore the mean *stellar* abundance is

$$\langle Z_\star \rangle = \frac{{}_0\int^\infty Z\, dM_\star(Z)}{\int_0^\infty dM_\star} = p\,, \tag{7.9}$$

therefore the mean metallicity equals the yield, p.

The G-dwarf problem
For the MW, $M_g/M \sim 0.1$ The abundance of star forming now is therefore $Z(M_g = 0.1\,M) = -p\log(0.1) \approx p/2.3$ If this is to be close to the solar abundance, $Z_\odot \approx 0.02$, then $p \approx 0.009$. Then we find that the fraction of stars with $Z \leq Z_\odot/4$ is $M_\star(\leq Z_\odot/4) = 0.44\,M$, or nearly half of MW stars should have abundance $\leq Z_\odot/4$. The observed value is nearer to 2 per cent. This is called the G-dwarf problem, because it was first encountered in G-dwarfs. Basically our theory predicts there should be many more low Z stars than are observed. The solutions is probably that the MW is far from a closed box, and has been losing a large fraction of its mass.

Chapter 8

Plasma effects

- Maxwell's equations and their units
- plane waves in empty space
- plane waves in a plasma: conductivity, dielectric constant, dispersion, plasma frequency, phase and group velocity, pulsar dispersion measurements
- polarisation
- Faraday rotation

Further reading: *Rybicki & Lightman, Chapter 2 and paragraphs 8.1 and 8.2*

8.1 Maxwell's equations

In Gaussian units, Maxwell's equations (ME) that relate the electric field ($\mathbf{E}$) and magnetic field ($\mathbf{B}$) to the charge density ρ and current ($\mathbf{J}$) are

$$\nabla \mathbf{E} = 4\pi\rho \tag{8.1}$$

$$\nabla \mathbf{B} = 0 \tag{8.2}$$

$$\nabla \times \mathbf{E} + \frac{1}{c}\frac{\partial \mathbf{B}}{\partial t} = 0 \tag{8.3}$$

$$\nabla \times \mathbf{B} = \frac{4\pi}{c}\mathbf{J} + \frac{1}{c}\frac{\partial \mathbf{E}}{\partial t}\,. \tag{8.4}$$

In these units, the electron charge $e = 4.8 \times 10^{-10}$ esu. The Lorentz force on a charge q moving with speed $\mathbf{v}$ is

$$\mathbf{F} = q\left(\mathbf{E} + \frac{\mathbf{v}}{c} \times \mathbf{B}\right). \tag{8.5}$$

See the appendix of Jackson's *Classical Electrodynamics* for a good introduction to the various choices of units of the ME.

8.1.1 Plane waves

Plane waves are solutions to Maxwell's equations. Write the $\mathbf{E}$ component of the wave as $\mathbf{E}(\mathbf{r}, t) = E_0\,\hat{\mathbf{a}}_0 \exp\left(i\left(\mathbf{kr} - \omega t\right)\right)$ (and similarly for $\mathbf{B}$), with the understanding that the electric field is the *real* part of this complex expression.

(a) in empty space: $\rho = \mathbf{J} = 0$
Substituting the *Ansatz*

$$\begin{aligned} \mathbf{E} &= E_0\,\hat{\mathbf{a}}_0 \exp\left(i\left(\mathbf{k}\cdot\mathbf{r} - \omega t\right)\right) && (8.6)\\ \mathbf{B} &= B_0\,\hat{\mathbf{a}}_1, \exp\left(i\left(\mathbf{k}\cdot\mathbf{r} - \omega t\right)\right), && (8.7) \end{aligned}$$

where E_0 and B_0, and $\hat{\mathbf{a}}_{0,1}$ are all constants, into the ME yields the algebraic relations

$$\begin{aligned} i\,\mathbf{k}\cdot\mathbf{E} &= 0 && (8.8)\\ i\,\mathbf{k}\cdot\mathbf{B} &= 0 && (8.9)\\ i\,\mathbf{k}\times\mathbf{E} - \frac{i\omega}{c}\,\mathbf{B} &= 0 && (8.10)\\ i\,\mathbf{k}\times\mathbf{B} + \frac{i\omega}{c}\mathbf{E} &= 0\,. && (8.11) \end{aligned}$$

These equations require $(\hat{\mathbf{a}}_0\,\hat{\mathbf{a}}_1, \mathbf{k})$ to be a *right-handed triad*, the amplitudes

$$E_0 = B_0\,, \tag{8.12}$$

and speed of the waves given by the dispersion relation

$$k^2\,c^2 = \omega^2\,. \tag{8.13}$$

As expected, this solution is a transverse wave moving with speed c, the speed of light. Since the ME are linear, any linear combination of such waves

will be a solution as well. In particular star light will be a linear combination of many such *monochromatic waves*. The amount of energy carried by the wave follows from application of the *Poynting* vector $\mathbf{S} = (c/4\pi)\,\mathbf{E} \times \mathbf{B}$: hence the energy flux $F \propto |\mathbf{S}| \propto E_0^2$.

(b) in a plasma
Making the same *Ansatz* in the more general case yields

$$
\begin{aligned}
i\,\mathbf{k}\cdot\mathbf{E} &= 4\pi\,\rho && (8.14)\\
i\,\mathbf{k}\cdot\mathbf{B} &= 0 && (8.15)\\
i\,\mathbf{k}\times\mathbf{E} - \frac{i\omega}{c}\,\mathbf{B} &= 0 && (8.16)\\
i\,\mathbf{k}\times\mathbf{B} + \frac{i\omega}{c}\mathbf{E} &= \frac{4\pi}{c}\,\mathbf{J}\,. && (8.17)
\end{aligned}
$$

We will consider a globally neutral *plasma* with electron density n, such that the charge density[1] $\rho = -n\,e$, and the current associated with moving electrons is $\mathbf{J} = -ne\,\mathbf{v}$. The solution to the equation of motion for the electron,

$$m\,\dot{\mathbf{v}} = -e\,\mathbf{E} \tag{8.18}$$

if we neglect the $\mathbf{B}$ term (which is smaller by $\sim v/c$) is[2]

$$\mathbf{v} = \frac{e\,\mathrm{E}}{i\,\omega\,m}\,, \tag{8.19}$$

and hence the current density defines the *conductivity*, σ, as

$$
\begin{aligned}
\mathbf{j} &= \sigma\,\mathbf{E} && (8.20)\\
\sigma &= \frac{i\,n\,e^2}{\omega\,m}\,. && (8.21)
\end{aligned}
$$

Substitution into the equation for charge conservation, $\dot{\rho} + \nabla\mathbf{J} = 0$, yields $\rho = \sigma\omega^{-1}\,\mathbf{k}\cdot\mathbf{E} = 0$. Application of the ME. $i\,\mathbf{k}\cdot\mathbf{E} = 4\pi\rho$, yields $i\,\epsilon\,\mathbf{k}\cdot\mathbf{E} = 0$ where the *dielectric constant* ϵ, is

$$\epsilon = 1 - \frac{4\pi\sigma}{i\,\omega}\,. \tag{8.22}$$

[1] Why are we neglecting the protons?
[2] Demonstrate as an exercise.

In terms of σ and ϵ, the ME now read

$$\begin{aligned} i\,\epsilon\,\mathbf{k}\cdot\mathbf{E} &= 0 && (8.23)\\ i\,\mathbf{k}\cdot\mathbf{B} &= 0 && (8.24)\\ i\,\mathbf{k}\times\mathbf{E} - \frac{i\omega}{c}\,\mathbf{B} &= 0 && (8.25)\\ i\,\mathbf{k}\times\mathbf{B} + \frac{i\omega}{c}\epsilon\mathbf{E} &= 0\,, && (8.26) \end{aligned}$$

similar to those in empty space apart from the appearance of ϵ. As before, these describe a transverse EM wave, but the dispersion relation is modified to

$$c^2\,k^2 = \epsilon\,\omega^2\,, \tag{8.27}$$

which shows that the speed of the wave now depends on ω, *i.e.* the plasma is *dispersive*. In terms of the *plasma frequency*, ω_p,

$$\omega_p^2 \equiv \frac{4\pi n\,e^2}{m} \tag{8.28}$$

the dispersion relation is

$$c^2k^2 = \omega^2 - \omega_p^2\,. \tag{8.29}$$

Waves with $\omega < \omega_p$ have imaginary $k = ik_I$, hence are exponentially damped, $\mathbf{E} \propto \exp(-k_I\,r)$. Presence of electron charge introduces a *plasma cut-off frequency* ω_p, below which the plasma does not allow radiation to propagate. This causes the reflection of ~ 1 MHz EM waves off the earth's ionosphere for example.

8.1.2 Group and phase velocity

The phase-velocity of a wave with $\omega = \omega(k)$ is

$$v_{\rm ph} \equiv \frac{\omega}{k} = \frac{c}{n_r}\,, \tag{8.30}$$

whereas the group velocity

$$v_{\rm g} \equiv \frac{\partial\omega}{\partial k} = c\left(1 - \frac{\omega_p^2}{\omega^2}\right)^{1/2}\,. \tag{8.31}$$

In empty space these are the same, but in a plasma $v_{\rm g} < c$: this is the speed with which the energy of the EM wave moves. The index of refraction $n_r = \sqrt{\epsilon}$.

Application: pulsar timing measurements
A pulsar produces pulses of EM radiation with a range of frequencies ω. Since the speed of each of its constituent waves depends on ω, the times, t_p, the various waves arrive on earth willl depend on ω, the distance d to the pulsar, and on the properties of the ISM along the line of sight:

$$t_p(\omega) = \int_0^d \frac{ds}{v_{\rm g}(s)} \,. \tag{8.32}$$

For waves with $\omega \gg \omega_p$, $v_{\rm g}^{-1} \approx c^{-1}\,\left(1 + \omega_p^2/2\omega^2\right)$, hence

$$t_p(\omega) \approx \frac{d}{c} + \frac{4\pi e^2}{2\omega^2\, c\, m}\,\int_0^d n(s)\, ds \,. \tag{8.33}$$

The first term is simply the time the EM would take in empty space, the second depends on frequency. We can measure how the arrival time depends on ω, ${\rm d}t_p/{\rm d}\omega$, from which we can constrain the electron density to the pulsar, or if we know n, the pulsar's distance d:

$$\frac{{\rm d}t_p}{{\rm d}\omega} = -\frac{4\pi\, e^2}{\omega^3\, c\, m}\,\int_0^d n(s)\, ds \,. \tag{8.34}$$

8.2 Polarisation

8.2.1 Definitions and Stokes parameters

Light is said to be *polarised* if the plane in which its **E** vector lies (and hence also **B** $\perp$ **E**) is constrained, for example to a plane for *linearly polarised* light. Consider the sum of two plane waves,

$$\mathbf{E} = \mathbf{E}_1 + \mathbf{E}_2 \,, \tag{8.35}$$

where the $\mathbf{E}_i$s are both monochromatic, $\propto \exp(-i\omega t)$, and complex (with the convention that the physical electric field is the real part of the expression). Assume the wave travels along $\hat{z}$, then without loss of generality we have

$$\mathbf{E}_1 \;=\; E_1\, \exp\left(i(k\, z - \omega t)\right)\, \hat{x} \tag{8.36}$$

$$E_1 \;=\; \epsilon_1\, \exp(i\varphi_1) \,, \tag{8.37}$$

and similarly for $\mathbf{E}_2 \parallel \hat{y}$. ϵ is the real amplitude of the wave, φ its phase. For simplicity we can analyse the EM field at $\mathbf{r} = \mathbf{0}$.

(a): no phase difference
If $\varphi_1 - \varphi_2 = 0$, the physical field is $\mathbf{E} = \epsilon_1 \cos(\omega t - \varphi)\,\hat{\mathbf{x}} + \epsilon_2 \cos(\omega t - \varphi)\,\hat{\mathbf{y}}$, which is a harmonically varying vector, moving in a plane under an angle β, $\tan(\beta) = \epsilon_2/\epsilon_1$ with the $\hat{\mathbf{x}}$ direction. This is a *linearly polarised* ray.

(b): phase difference $= \pm\pi/2$
Assuming for simplicity that $\varphi_1 = 0$, and $\varphi_2 = \pm\pi/2$, the physical field has

$$\begin{aligned} \mathbf{E} &= \mathcal{R}\left[(\epsilon_1\,\hat{\mathbf{x}} \pm i\,\epsilon_2\hat{\mathbf{y}})\exp(-i\omega t)\right] && (8.38)\\ &= \epsilon_1\,\cos(\omega t)\,\hat{\mathbf{x}} \pm \epsilon_2\,\sin(\omega t)\,\hat{\mathbf{y}}\,. && (8.39)\end{aligned}$$

Such a ray is called *elliptically polarised*, because the $\mathbf{E}$ vector traces-out an ellipse with major/minor axis $\epsilon_{1,2}$. It is traced in anti-clockwise (clockwise) for $\phi_2 = \pi/2$ $(-\pi/2)$: these are called right, respectively left-hand polarised. The special case of $\epsilon_1 = \epsilon_2$ is called *circularly polarised*. In the more general case where $\varphi_1 \neq 0$, the EM -wave is also elliptically polarised, but the major/minor axes are not along our chosen $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ Cartesian axes.

The polarisation state of radiation is fully characterised by the following four *Stokes* parameters

$$\begin{aligned} I &\equiv \epsilon_1^2 + \epsilon_2^2 \quad \text{'intensity'} && (8.40)\\ Q &\equiv \epsilon_1^2 - \epsilon_2^2 \quad \text{'circularity'} && (8.41)\\ U &\equiv 2\,\epsilon_1\,\epsilon_2\,\cos(\varphi_1 - \varphi_2) && (8.42)\\ V &\equiv 2\,\epsilon_1\,\epsilon_2\,\sin(\varphi_1 - \varphi_2)\,. && (8.43)\end{aligned}$$

Given these definitions, $I^2 = Q^2 + U^2 + V^2$, which expresses the fact that monochromatic waves are always completely polarised. Most star light is not monochromatic and in that case

$$\frac{Q^2 + U^2 + V^2}{I^2} \leq 1 \qquad (8.44)$$

is called the *degree of polarization*. Reflected light tends to be polarised, hence the degree of polarisation can be used to judged whether light has been scattered. For example light scattered off reflection nebulae is polarised.

8.2.2 Application: Faraday rotation

Consider a circularly polarised beam of light moving parallel to a constant magnetic field:

$$\mathbf{E} = E_0 \left(\hat{\mathbf{x}} \mp i\hat{\mathbf{y}}\right) \exp(-i\omega t) \tag{8.45}$$

$$\mathbf{B}_0 = B_0 \,\hat{\mathbf{z}}\,. \tag{8.46}$$

The presence of the constant magnetic field introduces the *cyclotron frequency*, $\omega_B = eB_0/m\,c$. Neglecting the magnetic field associated with the EM wave compared to the external one, yields the equation of motion for the electrons:

$$m\dot{\mathbf{v}} = -e\,\mathbf{E} - \frac{e}{c}\,\mathbf{v}\times\mathbf{B}_0\,, \tag{8.47}$$

with solution[3]

$$\mathbf{v} = \frac{-i\,e}{m\left(\omega \pm \omega_B\right)}\,\mathbf{E}(t)\,. \tag{8.48}$$

Compare to the previous case with $B_0 = 0$, Eq. (8.19). We can follow the derivation in that section to find that the dielectric constant is now given by

$$\epsilon = 1 - \frac{\omega_p^2}{\omega\left(\omega \pm \omega_B\right)}\,, \tag{8.49}$$

and is *different* for the two different polarisations (left and right-handed). This means that the speed with which the polarised waves move, is different for right and left-handed polarisations. A corollary is that the plane in which linearly polarised light moves, will rotate as a function of position along the line of sight, a phenomenon called *Faraday rotation*.

The angle over which the plane rotates can be found by realising that the phase of the wave is given by $\varphi_{\mathrm{R,L}} = \int_0^d k_{\mathrm{R,L}}\,ds$, where $k_{\mathrm{R,L}} = (\omega/c)\,\sqrt{\epsilon_{\mathrm{R,L}}}$. The angle $\Delta\theta$ is just half of the difference in phase between R and L-handed polarisation. Hence

$$\Delta\theta = \frac{1}{2}\int_0^d (k_{\mathrm{R}} - k_{\mathrm{L}})\,ds \tag{8.50}$$

$$\approx \frac{2\pi e^3}{m^2\,c^2\,\omega^2}\int_0^d n(s)\,B_0(s)\,ds\,, \tag{8.51}$$

[3] Check!

where the approximation assumes $\omega \gg \omega_p$ and $\omega \gg \omega_B$. It depends on the integral of the product of electron density and magnetic field strength parallel to the line of sight. We can constrain this integral by measuring $\partial\Delta\theta/\partial\omega$, *i.e.* how the plane of polarisation depends on frequency. This provides us with a measure of the mean magnetic field along the line of sight if n is known, for example from pulsar timing measurements.

8.3 Exercises

1. The spectrum of EM radiation received from a pulsar has arrival times that depend on frequency as $\mathrm{d}t_p/\mathrm{d}\omega = 1.1 \times 10^{-5}$ s^2, and a Faraday rotation measure $\partial\Delta\theta/\partial\omega = 1.9 \times 10^{-4}$ rad s^{-1}, both at $\omega = 10^8$ Hz. Determine the mean magnetic field strength B_0 along the line of sight. (From Rybicky & Lightman, Exercise 8.3)

Appendix A

Statistics and basic conceps

A.1 Black Body radiation

The intensity for a Black Body of temperature T depends on frequency as

$$B_\nu(T) = \frac{2h\nu^3/c^2}{\exp(h\nu/kT) - 1}\,. \tag{A.1}$$

A.2 Maxwell distribution

The fraction of particles with velocity v in a Maxwellian distribution at temperature T is

$$f(v) = 4\pi \left(\frac{m}{2\pi kT}\right)^{3/2} v^2 \exp(-mv^2/2kT)\,. \tag{A.2}$$

A.2.1 Detailed balance

Consider a reaction

$$A + B \to C + D \tag{A.3}$$

and its inverse

$$C + D \to A + B\,. \tag{A.4}$$

In a system in equilibrium these reactions occur at the same rate. Since the photon distribution as well a the thermal distributions are known in equilibrium (Black Body and Maxwell distributions, respectively), we should be able to compute how the reaction rates for each of the above depends on

frequency of the radiation and velocity of the colliding particles. Since these dependencies are characteristic of the particles $A \cdots D$, the actual relations themselves do no require equilibrium. Such *detailed balance* considerations are hence a powerful tool to relate rate coefficients.

A.3 Boltzmann distribution

(*Rybicky & Lightman §9.5*)
Consider a system which can be in a variety of states, $i = 1, \cdots, N$, for example the energy levels n in a hydrogen atom. The *Boltzmann distribution* characterises the fraction of particles in a given state,

$$\frac{n_i}{n} = \frac{g_i \, \exp(-E_i/kT)}{Z} \,, \tag{A.5}$$

where g_i is the statistical weight of state of state i (*e.g.* $g_i = 2\sum_0^{n-1}(2l + 1) = 2n^2$ for the hydrogen atom), T the temperature, and E_i the energy above the ground state $n = 1$. Since $\sum n_i = n$, the *partition function* $Z = \sum_i g_i \, \exp(-E_i/kT)$.

A.4 Saha equation

(*Rybicky & Lightman §9.5*)
The Boltzmann equation applied to a thermal gas of ions and neutrals gives for the fraction of ions of which the free electron has a given velocity v, $dn_0^+(v)$ compared to neutrals n_0

$$\frac{\mathrm{d}n_0^+}{n_0} = \frac{g}{g_0} \, \exp(-\frac{\chi_I + mv^2/2}{kT}) \,, \tag{A.6}$$

where χ is the ionisation potential, g_0 is the statistical weight of the neutrals, and $g = g_0^+ \, g_e$ the product of the statistical weights of ions and electrons.

Treating the electrons as free particles in a box with volume $V = L_x L_y L_z$, the allowed quantum numbers in say the x-direction follows from requiring there to be an integer number of de Broglie wavelengths in L_x: $n_x \lambda_x = n_x(2\pi/k_x) = L_X$, hence $n_x = L_x \, p_x/h$ in terms of the electron's x momentum p_x. Taking into account the two possible spin states, the total number of

states in 3D is then $N = 2n_x n_y n_z = 2(L_x L_y L_z) d^3p/h^3$. The total number of electrons in V is $N_e = n_e V$ in terms of the electron number density n_e. Hence the density of states is $g_e = N/N_e = 2d^3p/n_e\, h^3$.

Inserting this in the Boltzmann equation yields

$$\frac{\mathrm{d}n_0^+}{n_0} = \frac{g_0^+}{g_0}\,\frac{8\pi m_e^3}{n_e\, h^3}\,\exp(-\frac{\chi_I + mv^2/2}{kT})\,v^2\,dv\,, \tag{A.7}$$

and integrated over all v it yields the *Saha equation* for the ratio of ionised over neutrals:

$$\frac{n_0^+\, n_e}{n_0} = \frac{2g_0^+}{n_0}\,\left(\frac{2\pi m_e kT}{h^2}\right)^{3/2}\,exp(-\frac{\chi_I}{kT})\,. \tag{A.8}$$

A.5 Einstein relations

Rybicky & Lightman, §1.6
Consider a two-level atom. Absorption of a photon may cause an excitation $1 \to 2$, and de-excitation may occur through spontaneous and induced emission. These processes are described by the *Einstein relations*. Denoting the number density of excited atoms by n_2 and ground state atoms by n_1, the rate equation is

$$\frac{\mathrm{d}n_2}{\mathrm{d}t} = B_{12}J_\nu, n_1 - A_{21}\, n_2 - B_{21}\, J_\nu\, n_2\,, \tag{A.9}$$

which describe absorption, spontaneous emission, and stimulated emission, respectively, in the presence of a radiation field with mean intensity J_ν.

When the radiation field is that of a Black Body, $J_\nu = B_\nu$, the system will get into an equilibrium with $\mathrm{d}n_2/\mathrm{d}t = 0$. The ratio n_2/n_1 is then given by the Boltzmann factor,

$$\frac{n_1}{n_2} = \frac{g_1}{g_2}\,\exp(-h\nu_{12}/kT)\,, \tag{A.10}$$

where g_i denotes the statistical weight of state i. Combing this with the expression of $B_\nu(T)$ for a Black Body yields the following 'Einstein relations' between the rate coefficients:

$$g_1 B_{12} = g_2 B_{21} \tag{A.11}$$

$$A_{21} = B_{21} \frac{2h\nu_{12}^3}{c^2} . \tag{A.12}$$

These atomic relations are always valid, not just in equilibrium.

Appendix B

Hints to the exercises

- Chapter 1

  1. Absorption with given optical depth τ reduces the flux from I_e to I_o, where $I_o/I_e = \exp(-\tau)$; hence the change in apparent magnitude $\Delta m = -2.5 \log_{10}(I_o/I_e) = -2.5 \log_{10}(\exp(-\tau)) \approx 1.086\,\tau$
  2. The optical depth after traversing a distance R due the grains is $\tau = n_d\,(\pi r_d^2)\,R$. Consider a small volume V, which contains a mass M_g of gas, and M_d of dust. The dust-to-gas ratio within this volume
  $$\psi \equiv \frac{M_d}{M_g} = \frac{n_d\,m_d\,V}{m_H\,n_H\,V} = \frac{n_d\,m_d}{m_H\,n_H}\,, \tag{B.1}$$
  hence the dust number density is $n_d = \psi\,(n_H\,m_H/m_d)$. Combining these equations yields
  $$\tau = \psi\,\frac{n_H\,m_h}{m_d}\,(\pi r_d^2)\,R\,. \tag{B.2}$$
  The mass m_d of a single dust grain is $m_d = (4\pi/3)\,\rho_d\,r_d^3$. Substituting the numbers gives $\tau = [0.04, 0.4]$ for $n_H = [10^2, 10^3]$ cm^{-3}.
  3. Assume $M_g = 10^{14}\,M_\odot$ and $R = 2$ Mpc for the cluster's gas mass, and radius. For uniform density (and assuming pure hydrogen gas, fully ionised), the electron density $n_e \approx 10^{-4}$ cm^{-3}. The Thomson optical depth to the centre is then $\tau = n_e\,\sigma_T\,R \approx 5 \times 10^{-4}$, where σ_T is the Thomson cross section.
  4. The Thomson optical depth is $\tau = 2 \times 10^{-6}$ using the equation above with the new numbers. The dust density assuming the

parameters from Exercise 2 is $n_d \approx 10^{-13}$ cm^{-3}. The dust optical depth is then $\tau_d = n_d\,(\pi r^2)\,R = 4 \times 10^{-4}$.

5. The infinitesimal optical depth due to Thomson scattering over a physical distance dl is $d\tau = n_e(a)\,\sigma_T\,dl$, where the physical electron density at expansion factor a depends on the value now at $a = 1$ as $n_e(a) = n_e(a = 1)a^{-3}$. In an expanding Universe, the relation between dl and da follows from the Hubble law:

$$dl = c\,dt = c\frac{dt}{da}\,da = \frac{c\,da}{a\,H(a)}\,, \tag{B.3}$$

and the Hubble constant at expansion factor a in an Einstein-de Sitter ($\Omega_m = 1$) Universe is

$$H(a) = H_0\,a^{-3/2}\,, \tag{B.4}$$

where $H_0 \equiv H(a = 1)$. Combining the above yields

$$\begin{aligned} \tau &= \int_{a_r}^{1} \frac{\sigma_T\,c\,n_e(a=1)}{H_0}\,\frac{a^{-3}\,da}{a\,a^{-3/2}} \\ &= \frac{2}{3}\,\frac{\sigma_T\,c\,n_e(a=1)}{H_0}\,\left[a_r^{-3/2} - 1\right] \\ &= 0.04\,. \end{aligned} \tag{B.5}$$

The numerical value assumes a Hubble constant $h = 0.72$, a baryon fraction $\Omega_b = 0.04$ and neglects helium.

- Chapter 3

1. Solution see exercise 8.3 in R& L.

- Chapter 4

- Chapter 5

1. Ionisations are due to photons and collisions, so the ionisation rate is

$$\frac{\mathrm{d}n_{\mathrm{HII}}}{\mathrm{d}t} = \Gamma\,n_{\mathrm{HI}} + \Gamma_e\,n_{\mathrm{HI}}\,n_e\,, \tag{B.6}$$

whereas the recombination rate is

$$\frac{\mathrm{d}n_{\mathrm{HI}}}{\mathrm{d}t} = \alpha\,n_{\mathrm{HII}}\,n_e\,. \tag{B.7}$$

Define the ionised fraction (note difference from original question) $x = n_{\rm HII}/n$, with $n = n_{\rm HI} + n_{\rm HII}$ and taking into account that $n_e = n_{\rm HII}$ for a pure hydrogen gas, then the equilibrium x is the solution to the quadratic equation

$$(\alpha + \Gamma_e)\, n\, x^2 + (\Gamma - \Gamma_e n)x - \Gamma = 0\,. \tag{B.8}$$

At low density, $\Gamma \gg \Gamma_e n$, $x \approx 1 - (\alpha + \Gamma_e)n/\Gamma$, whereas at high density, $x \approx \Gamma_e/(\alpha + \Gamma_e)$ if collisional ionisations are important, and $x \approx (\Gamma/\alpha n)^{1/2}$ if they are not.

2. The evolution of the neutral density due to recombinations is

$$\frac{{\rm d}n_{\rm HI}}{{\rm d}t} = \alpha\, n_{\rm HII}\, n_e\,, \tag{B.9}$$

or in terms of $x \equiv n_{\rm HII}/n = n_e/n$, $1 - x = n_{\rm HI}/n$

$$\frac{\rm d}{{\rm d}t}(1 - x) = \alpha\, x^2\, n\,. \tag{B.10}$$

Defining the dimensionless time $\tau \equiv t/(\alpha\, n)$, this becomes

$$\frac{\rm d}{{\rm d}\tau}(1 - x) = x^2\,, \tag{B.11}$$

which has the solution

$$x = \frac{1}{1 + \tau}\,, \tag{B.12}$$

which has the correct initial condition $x = 1$ at $t = 0$. For $t = t_r = 1/(\alpha n)$, $\tau = 1$, and the ionised fraction is $x = 1/2$, as is the neutral fraction.

3. The formation rate is

$$\frac{{\rm d}n({\rm H_2})}{{\rm d}t} = \frac{1}{2}\,\epsilon\,(\pi a^2)\, n_d\, n_{\rm H}\, v_{\rm H}\,, \tag{B.13}$$

where ϵ is the dimensionless sticking coefficient of Hydrogen atoms on a grain, (πa^2) is the geometric cross section of a grain (dimensions of a surface area), n_d and $n_{\rm H}$ are the number densities of grains and Hydrogen atoms, and $v_{\rm H}$ is mean relative velocity of

grains and atoms. The factor 1/2 arises since two hydrogen atoms need to be adsorbed for the reaction to molecular hydrogen to take place. The destruction rate is

$$\begin{aligned} \frac{\mathrm{d}n(\mathrm{H_2})}{\mathrm{d}t} &= -n(\mathrm{H_2}) \int_{\nu_{\rm th}}^{\infty} \frac{4\pi J(\nu)}{h\nu}\, \sigma(\nu)\, d\nu\,, \\ &\equiv n(\mathrm{H_2})\, \Gamma\,, \end{aligned} \tag{B.14}$$

where $n(\mathrm{H_2})$ is the number density of molecular hydrogen, $J(\nu)$ is the mean intensity, $[4\pi J(\nu)/h\nu] = \mathrm{cm^{-2}\ s^{-1}\ Hz^{-1}}$, $\sigma(\nu)$ is the cross section for photo-destruction of molecular hydrogen due to photons with frequency ν (units of area), and ν is the frequency. The integral starts at the minimum frequency $\nu_{\rm th}$ where a photon can destroy molecular hydrogen. In equilibrium the rates balance, and with $x = n(\mathrm{H_2})/n$, where $n \equiv n(\mathrm{H_2}) + n_{\rm H}$, the equilibrium molecular fraction x is

$$x = \frac{1}{1 + \Gamma/k\, n_d}\,, \tag{B.15}$$

where the rate constant $k \equiv (1/2)\epsilon\, (\pi a^2)\, n_d\, v_{\rm H}$ with dimension $\mathrm{cm^{-3}\ s^{-1}}$.

- Chapter 6

  1. See R&L, exercise 10.6
  2. C, O and Si have 6, 8 and 14 protons, respectively. Therefore C IV=C^{3+} and O VI=O^{5+} have both 3 electrons, and their ground state is $1s^2 2s^1$; whereas Si IV=Si^{3+} has 11 electrons, hence its ground state is $1s^2 2s^2 2p^6 3s^1$. Neglecting filled shells, their valence electrons comprise a single s electron, similar to the Alkali metals, Na, K, Rb, etc. The doublet structure has the same origin as that of the famous 'Sodium D-lines'. The ground state has $l = 0$, $s = 1/2$ and hence the term $^2S_{1/2}$, the excited p-state has $l = 1$ and $s = 1/2$ hence there are two possibilities, $^2P_{1/2}$ and $^2P_{3/2}$. Applying Landés rule predicts that the $^2P_{1/2}$ state is more bound than the $^2P_{3/2}$ state, so the energy levels from most to least bound are $^2S_{1/2}$, $^2P_{1/2}$ and $^2P_{3/2}$. Transitions from the excited P states to the ground S state have $\Delta m = 0\,, \pm 1$ and $\Delta l = 1$ so are

dipole allowed. The degeneracy of the P states, $2J+1$, are 2 for $^2P_{1/2}$ and 4 for $^2P_{3/2}$, ratio 1:2, hence the higher energy transition $^2P_{3/2} \rightarrow {}^2S_{1/2}$ is twice as strong as the lower energy component of the doublet, $^2P_{1/2} \rightarrow {}^2S_{1/2}$.

- Chapter 7

  1. Let $\tau \equiv t/t_r$, then

$$\begin{aligned} \frac{\mathrm{d}R}{\mathrm{d}\tau} &= \frac{1}{3} R_s (1-\exp(-\tau))^{-2/3} \exp(-\tau) \\ &= \frac{1}{3} R_s \left(1-\exp(-\tau)\right)^{-2/3} - \frac{1}{3} R_s \left(1-\exp(-\tau)\right)^{1/3} \\ &= \frac{R_s^3}{3R^2} - \frac{R}{3} \,. \end{aligned} \tag{B.16}$$

     Comparing to Eq. (4.1),

$$\frac{\mathrm{d}R}{\mathrm{d}t} = \frac{\dot{N}_\gamma}{4\pi R^2 \alpha\, n^2} - \frac{R}{3} \,, \tag{B.17}$$

     shows that these are the same, provided

$$R_s^3 = \frac{\dot{N}_\gamma}{(4\pi/3)\,\alpha\, n^2} \,, \tag{B.18}$$

     which is indeed the expression for the Strömgren radius.

  2. This exercise is solved in Dyson & Williams on p. 71

Commonly used symbols and results

Flux F, $[F]$= erg s^{-1} cm^{-2}, $[F_\nu]$ = erg s^{-1} cm^{-2} Hz^{-1}

Intensity I, $[I_\nu]$ = erg s^{-1} cm^{-2} ster^{-1} Hz^{-1}

Mean intensity $J_\nu = \frac{1}{4\pi} \int I_\nu \, d\Omega$

Emission coefficient j_ν defines the energy produced per unit volume, $dE_\nu = j_\nu \, dV \, d\Omega \, dt \, d\nu$, $[j_\nu]$=erg cm^{-3} ster^{-1} s^{-1} Hz^{-1}

Absorption coefficient α_ν and optical depth τ_ν, $dI_\nu = -\alpha_\nu \, I_\nu \, ds = -I_\nu \, d\tau_\nu$

Optical depth due to solid spheres, $d\tau_\nu = Q(\nu) \, n \, (\pi \, r^2) \, ds$, with $Q(\nu)$ the frequency-dependent extinction efficiency

Plasma frequency $\omega_p^2 \equiv 4\pi n \, e^2/m$ and plasma dispersion relation $c^2 k^2 = \omega^2 - \omega_p^2$

Metallicity $Z = M_z/M_g$

General collisional reaction rate for A+B$\rightarrow$ C: $\frac{\mathrm{d}}{\mathrm{d}t} n_\mathrm{C} = \sigma \, n_\mathrm{A} \, n_\mathrm{B} \, \langle v_\mathrm{AB} \rangle \equiv k \, n_\mathrm{A} \, n_\mathrm{B}$.

Particular case of photo-ionisation: $\frac{\mathrm{d}}{\mathrm{d}t} n_\mathrm{HII} = n_\mathrm{HI} \int_{\nu_\mathrm{th}}^{\infty} \sigma(\nu) \, \frac{4\pi J(\nu)}{h\nu} \, d\nu \equiv n_\mathrm{HI} \, \Gamma$.

Recombination: $\frac{\mathrm{d}}{\mathrm{d}t} n_\mathrm{HI} = \alpha \, n_\mathrm{HII} \, n_\mathrm{e}$

Absorption by grains: $\alpha_\nu = \frac{\mathrm{d}\tau_\nu}{\mathrm{d}s} = n_d\,(\pi a^2)\,Q_\nu$

Associated extinction $A_\nu = (2.5 \log e)\,\tau_\nu$

Kirchoff's law applied to dust with temperature T_d: $j_\nu = \alpha_\nu\,B_\nu(T_d)$

Notation for terms: ${}^{2S+1}L_J$

Hund's rules:

- terms with larger spin S tend to lie lower in energy
- for a configuration with given S, terms with larger L tend to lie lower in energy

Selection rules for dipole radiation: $\Delta l = \pm 1$, $\Delta m = 0, \pm 1$, $\Delta s = 0$

Photo-Ionisation equilibrium: $n_{\mathrm{HI}}\Gamma = \alpha n_{\mathrm{HII}} n_e$

Case B recombination rate $\alpha_B = \sum_{n=2}^{\infty} \alpha_n = \alpha_A - \alpha_1$.

Photo-heating rate $\mathrm{H} = n_{\mathrm{HI}} \int_{\nu_{\mathrm{th}}}^{\infty} \frac{4\pi J_\nu}{h\nu}\,(h\nu - h\nu_{\mathrm{th}})\,\sigma_\nu\,d\nu$

Recombination cooling rate $\mathrm{L} = \alpha_B\,n_{\mathrm{HII}}\,n_e\,\eta\,kT$ with $\eta \approx 1$

Line cooling due to collisionally excited lines: $\mathrm{L} = n_e n_1 q_{12} \Delta E_{12}$ with $q_{12}(T) \propto T^{-1/2} \exp(-\Delta E_{12}/kT)$

Contents

Bibliography

Abel T., Wandelt B. D., 2002, MNRAS, 330, L53

Altay G., Croft R. A. C., Pelupessy I., 2008, MNRAS, 386, 1931

Altay G., Theuns T., 2013, MNRAS, 434, 748

Camps P., Trayford J. W., Baes M., Theuns T., Schaller M., Schaye J., 2016, MNRAS, 462, 1057

Ferland G. J. et al., 2013, Rev. Mexicana Astron. Astrofis., 49, 137

Gnedin N. Y., Abel T., 2001, New A, 6, 437

Graziani L., Maselli A., Ciardi B., 2013, MNRAS, 431, 722

Hilborn R. C., 2002, ArXiv Physics e-prints

Matsuura M. et al., 2011, Science, 333, 1258

Mellema G., Iliev I. T., Alvarez M. A., Shapiro P. R., 2006, New A, 11, 374

Pawlik A. H., Schaye J., 2008, MNRAS, 389, 651

Rijkhorst E.-J., Plewa T., Dubey A., Mellema G., 2006, A&A, 452, 907

Rosdahl J., Blaizot J., Aubert D., Stranex T., Teyssier R., 2013, MNRAS, 436, 2188

Trayford J. W. et al., 2017, MNRAS, 470, 771

Wiersma R. P. C., Schaye J., Smith B. D., 2009a, MNRAS, 393, 99

Wiersma R. P. C., Schaye J., Theuns T., Dalla Vecchia C., Tornatore L., 2009b, MNRAS, 399, 574