ASTR705: The Interstellar Medium - Spring 2023

Prof. Loren Anderson
Phone: 304-293-4884; 304-508-8486
Email: loren.anderson@mail.wvu.edu
Meeting Times: 10:30 - 11:20 M W F, White G51
Office Hours: M 9:30-10:30, Th 12:00-1:00, and by appointment
Required Textbook: "Physics of the Interstellar and Intergalactic Medium," Draine, B.T.
Suggested: "Physics And Chemistry of the Interstellar Medium," Kwok, S.

Course Description
An in-depth look at the interstellar medium (ISM), the material in between stars, with a focus on our own Milky Way Galaxy. Topics covered include the composition of our Galaxy, the phases of the ISM, the properties of the gas and dust in the ISM, dust and gas chemistry, magnetic fields, dynamic processes, and star formation.

Expected Learning Outcomes
By the end of this course, you will be able to:
-describe the phases and composition of the ISM;
-analyze observational data from gas and dust in the ISM;
-evaluate under which circumstances certain dynamical effects will dominate interactions between the various ISM phases.

Suggested Prerequisite
ASTR601 "Astrophysics Seminar". Students who have not taken this course should study the modules on estimation, units, radiative transfer and blackbody emission.

Grading
The course will be graded out of 400 points, with 20 points for each of the 10 homework assignments, 50 points for the midterm, 100 points for the final exam, and 50 points for the oral presentation.

Grades will be assigned based on the points earned:
350+ = A
300+ = B
250+ = C
200+ = D
200− = F

Homework I will assign regular homework throughout the course. I will not accept late assignments, except in extraordinary circumstances. Expect to spend two hours on reading and homework for each hour of in-class time. Homework will generally be assigned on Fridays and due the following Friday; we will spend part of that class period reviewing solutions.

Midterm Exam Halfway through the course there will be a take-home midterm exam.

Final Exam There will be a single final, comprehensive, take-home exam given during the final class period and due before our specified finals time.

Topical Presentations Throughout the course I will allow each of you to present one small aspect of a current topic. These can take the form of journal-club-like reviews of classic papers in the field or black-board presentations, at your discretion. Full credit will be given for simply presenting.

Inclusivity Statement The West Virginia University community is committed to creating and fostering a positive learning and working environment based on open communication, mutual respect, and inclusion. If you are a person with a disability and anticipate needing any type of accommodation in order to participate in your classes, please advise your instructors and make appropriate arrangements with the Office of Accessibility Services. (https://accessibilityservices.wvu.edu/)

More information is available at the Division of Diversity, Equity, and Inclusion (https://diversity.wvu.edu/) as well. [adopted 2-11-2013]

Academic Integrity Statement The integrity of the classes offered by any academic institution solidifies the foundation of its mission and cannot be sacrificed to expediency, ignorance, or blatant fraud. Therefore, instructors will enforce rigorous standards of academic integrity in all aspects and assignments of their courses. For the detailed policy of West Virginia University regarding the definitions of acts considered to fall under academic dishonesty and possible ensuing sanctions, please see the West Virginia University Academic Standards Policy. Should you have any questions about possibly improper research citations or references, or any other activity that may be interpreted as an attempt at academic dishonesty, please see your instructor before the assignment is due to discuss the matter.

Other Resources There are many good books and resources available for the study of the ISM. I think the Draine book is the best, but these may be of interest as well. I have copies of many of these, if you want to borrow one at any point in the course.

- "Dust in the Galactic Environment: 2nd Edition" by Whittet, ISBN 0-7503-0624-6 (2003). A nice treatment of dust in the ISM.
- "An Introduction to Star Formation," by Ward-Thompson & Whitworth, ISBN 978-0-521-63030-6. Focuses on clouds and star formation.
- "Astrophysics of Gaseous Nebulae and Active Galactic Nuclei: 2nd Edition," by Osterbrock & Ferland, ISBN 1-891389-34-3 (2006). A classic but biased to optical observations
- "The Physics and Chemistry of the Interstellar Medium," by Tielens, ISBN 0- 521-82634-9 (2005). Similar to Kwok.
- "The Interstellar Medium," by Lequeux, ISBN 3-540-21326-0 (2003) – Good, but oddly

ordered.

- "Astrophysics of the Diffuse Universe," by Dopita & Sutherland, ISBN 3-540- 43362-7 (2003). Also good

- "The Physics of the Interstellar Medium: 2nd Edition" by Dyson & Williams, ISBN 0-7503-0460-X (1997) - older, thinner, not very deep, but a good quick read, especially for H II regions

- "Physical Processes in the Interstellar Medium," by Spitzer, ISBN 0-471-02232-2 (1978) – Older standard in the field, but dated and difficult to follow

- Harvard Astronomy ay201b Online Course: http://ay201b.wordpress.com/

- NRAO Essential Radio Astronomy Online Course: http://www.cv.nrao.edu/course/astr534/ERA.shtml

- Aaron Parson's YouTube channel: https://www.youtube.com/user/AaronRobertParsons/videos

- CalTech's wiki: https://casper.berkeley.edu/astrobaki/index.php/Main_Page

Day	Date	Topics	Reading	Assignments
Mon.	Jan. 9	Basics of the ISM and MW	Draine Ch. 1	
Wed.	Jan. 11	Basics of the ISM and MW	Draine Ch. 1	
Fri.	Jan. 13	Stat. Mech., Radiative Transfer	Kwok, Ch. 2,3	
Mon.	Jan. 16	MLK Day - No Class		
Wed.	Jan. 18	Stat. Mech., Radiative Transfer	Kwok, Ch. 2,3	
Fri.	Jan. 20	Stat. Mech., Radiative Transfer	Kwok, Ch. 2,3	HW1
Mon.	Jan. 23	Introduction to (Line) Emission	Notes; Draine Ch. 2,3,4,6,7,9,17	
Wed.	Jan. 25	Introduction to (Line) Emission	Notes; Draine Ch. 2,3,4,6,7,9,17	
Fri.	Jan. 27	Introduction to (Line) Emission	Notes; Draine Ch. 2,3,4,6,7,9,17	HW1 Due; HW2
Mon.	Jan. 30	Atomic Gas	Draine Ch. 8, 29, 30	
Wed.	Feb. 1	Atomic Gas	Draine Ch. 8, 29, 30	
Fri.	Feb. 3	Atomic Gas	Draine Ch. 8, 29, 30	HW2 Due; HW3
Mon.	Feb. 6	Atomic Gas	Draine Ch. 8, 29, 30	
Wed.	Feb. 8	Atomic Gas	Draine Ch. 8, 29, 30	
Fri.	Feb. 10	Atomic Gas	Draine Ch. 8, 29, 30	HW3 Due; HW4
Mon.	Feb. 13	Molecular Gas	Draine Ch. 5, 32, 32, 33; Kwok Ch. 7, 9	
Wed.	Feb. 15	Molecular Gas	Draine Ch. 5, 32, 32, 33; Kwok Ch. 7, 9	
Fri.	Feb. 17	Molecular Gas	Draine Ch. 5, 32, 32, 33; Kwok Ch. 7, 9	HW4 Due; HW5
Mon.	Feb. 20	Molecular Gas	Draine Ch. 5, 32, 32, 33; Kwok Ch. 7, 9	
Wed.	Feb. 22	Molecular Gas	Draine Ch. 5, 32, 32, 33; Kwok Ch. 7, 9	
Fri.	Feb. 24	Molecular Gas	Draine Ch. 5, 32, 32, 33; Kwok Ch. 7, 9	HW5 Due
Mon.	Feb. 27	Midterm Review	Midterm	
Wed.	Mar. 1	Continuous Radiation	Notes; Draine Ch. 2,3,4,6,7,9,17	
Fri.	Mar. 3	Midterm Review		
Mon.	Mar. 6	Dust	Draine Ch. 12.4, 21, 22, 23, 24	
Wed.	Mar. 8	Dust	Draine Ch. 12.4, 21, 22, 23, 24	
Fri.	Mar. 10	Dust	Draine Ch. 12.4, 21, 22, 23, 24	HW6
MWF	Mar. 13-17	Spring Recess - No Class		
Mon.	Mar. 20	Dust	Draine Ch. 12.4, 21, 22, 23, 24	
Wed.	Mar. 22	Dust	Draine Ch. 12.4, 21, 22, 23, 24	
Fri.	Mar. 24	Dust	Draine Ch. 12.4, 21, 22, 23, 24	HW6 Due; HW7
Mon.	Mar. 27	Ionized Gas	Draine Ch. 10-15, 18, 27, 34	
Wed.	Mar. 29	Ionized Gas	Draine Ch. 10-15, 18, 27, 34	
Fri.	Mar. 31	Ionized Gas	Draine Ch. 10-15, 18, 27, 34	HW7 Due; HW8
Mon.	Apr. 3	Ionized Gas	Draine Ch. 10-15, 18, 27, 34	
Wed.	Apr. 5	Ionized Gas	Draine Ch. 10-15, 18, 27, 34	
Fri.	Apr. 7	Ionized Gas	Draine Ch. 10-15, 18, 27, 34	
Mon.	Apr. 10	Spring Holiday - No Class		
Wed.	Apr. 12	Dr. Anderson Travelling		
Fri.	Apr. 14	ISM Dynamics and SF	Draine Ch. 36, 41, 42	HW8 Due; HW9
Mon.	Apr. 17	ISM Dynamics and SF	Draine Ch. 36, 41, 42	
Wed.	Apr. 19	ISM Dynamics and SF	Draine Ch. 36, 41, 42	
Fri.	Apr. 21	ISM Dynamics and Star Formation	Draine Ch. 36, 41, 42	HW9 Due; HW10
Mon.	Apr. 24	ISM Dynamics and Star Formation	Draine Ch. 36, 41, 42	
Wed.	Apr. 26	Makeup/Special Topic		
Fri.	Apr. 28	Final Review	HW10 Due; Final	
Th.	May 4	2-4pm	Final Due	

These notes serve as a summary of the material we will cover in ASTR705 this semester. My hope is that you find these a useful reference, in addition to our textbook.

Editor's note: text within square brackets represents notes to myself, since I am lecturing directly from this document!

Contents

Chapter 1

Introduction

1.1 Basics (Draine Ch. 1)

The interstellar medium, or ISM, is the material between stars. Stars form from material in the ISM, and when they die that material is recycled back into the ISM from which the next generation of stars will form. Whether you are studying the ISM directly or not, it shows up in all observations (including observations of pulsars!). About 10% of all the baryons in the Milky Way (MW) are in the ISM, but its influence is much greater than that implied by this number.

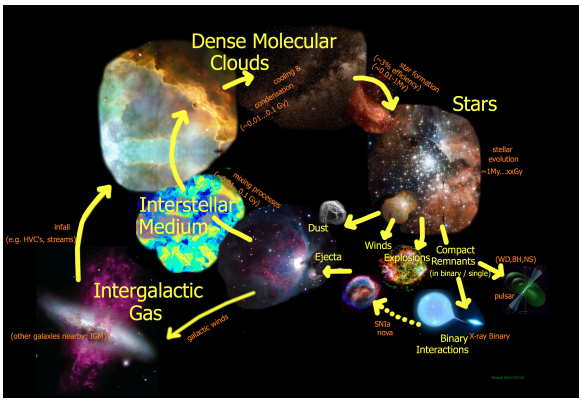


Figure 1.1: The cycle of material in the ISM. Stars are born of gas from the ISM and return these elements to the ISM when they die.

So, why should you care about the ISM?
(1) You probably need this class for credit;
(2) The ISM is one of the main constituents of spiral galaxies like our own;
(3) Material in the ISM forms stars, and so is critical to understanding galaxies;
(4) The physics of the ISM is really interesting, and touches on many different areas;
(5) It is probably a good time to admit that astronomy doesn't really need to have reasons, because it is cool enough already.

We want to understand the organization, distribution, physical conditions, and evolution of the ISM. We will focus on the ISM in our Galaxy (capital "G"!), since it is by far the best studied and understood. The topics extend naturally to the IGM, the intergalactic medium.

What is the ISM made of?

- Gas (atomic and molecular), including plasma (ions and electrons). This is the dominant form in terms of mass in the Galaxy.
- Dust. Small $\lesssim 1\mu$m particles mixed with the gas. These are molecules too big to name. About 1% of the gas mass is dust. Created by evolved stars, with small contribution from supernovae.
- Cosmic Rays (CRs). Ions and electrons with extremely high energies that zoom around the Galaxy. Because they are treated as individual particles, this phase is distinct from a plasma.
- Dark matter (DM). No EM interaction, only gravity for some reason. Non-barionic. We won't deal with this component explicitly.

And what fields are there?

- EM radiation. Primarily photons from stars, the CMB, gas (incl. plasmas), dust, and relativistic electrons (synchrotron).
- The magnetic field (guides ionized particles). The magnetic field strength is notoriously difficult to quantify, but can be very important.
- The gravitational field.

This course will cover as many of these as is possible during the term. When I took a class on the ISM as a gard student, it went on for three semesters and never got to star formation. The ISM is a very rich subject!

1.1.1 The ISM and the MW (not in book)

The MW is a spiral (disk) galaxy. A spiral galaxy has a bulge (old stars) and a disk (old+young stars). The corona, or halo is a spherical component surrounding the disk. Spiral (and irregular) galaxies have ongoing star formation. Stars are made from gas (+dust) and therefore spiral galaxies have gas+dust. Elliptical galaxies by-and-large only have stars. We hates them.

The stars and gas in the MW extend to a Galactocentric radius (distance from GC, R_G) of about 25 kpc. The Sun is at a R_G of ~ 8.5 kpc from the GC. The distribution of baryonic matter decreases with R_G. The percentage of DM increases with R_G.

Using R_G is great for understanding the MW, but our measurements are actually made in

Figure 1.2: Spiral galaxies NGC1300 (left), which displays a prominent bar, and M51 (right), which has a companion galaxy, but no bar. Note the obvious dust lanes in both galaxies and the blue colors indicative of ongoing (massive) star formation.

Figure 1.3: The elliptical galaxy M3. Note how boring and stupid it is due to the lack of dust, and the yellow color indicative of an older stellar population.

Galactic coordinates. These coordinates have as their origin the location of Sun. They use ℓ for Galactic longitude and b for Galactic latitude. b is defined as zero in the Galactic mid-plane, at positive values up to $+90^\circ$ toward the North Galactic pole, and negative values down to $+90^\circ$ toward the South Galactic pole. ℓ is defined increasing toward the east along the mid-plane, from $0/360^\circ$ at the GC to 180° in the Galactic anticenter. We define quadrants such that $0-90^\circ$ is quadrant I, $90-180^\circ$ is quadrant II $180-270^\circ$ is quadrant III $270-360^\circ$ is quadrant IV. This is a Sun-centered definition! We define the "inner Galaxy" as the region within the Sun's Galactocentric radius (and therefore orbit), and the "outer Galaxy" as the contrary. The inner Galaxy has by far the majority of stars, star formation. and ISM material. Quadrants I and IV are often referred to as the inner Galaxy.

[look at plates in book]

Within 15 kpc of the Galactic center, the total mass of the MW is $\sim 10^{11}\,M_\odot$. Of this, about half ($\sim 5\times10^{10}\,M_\odot$) is stars, about half ($\sim 5\times10^{10}\,M_\odot$) is DM, and a small percentage ($\sim 7\times10^{9}\,M_\odot$) is gas. Dust is only 1% that of gas. Notice the twiddles! These are all quite uncertain. Of the gas, $\sim 60\%$ is atomic hydrogen, $\sim 20\%$ is molecular hydrogen (H_2), and $\sim 20\%$ is ionized hydrogen. Helium adds $\sim 25\%$ to these numbers. (It is worth noting that $\sim 10\%$ He by number is $\sim 40\%$ by mass.) The above numbers differ for each type of Galaxy.

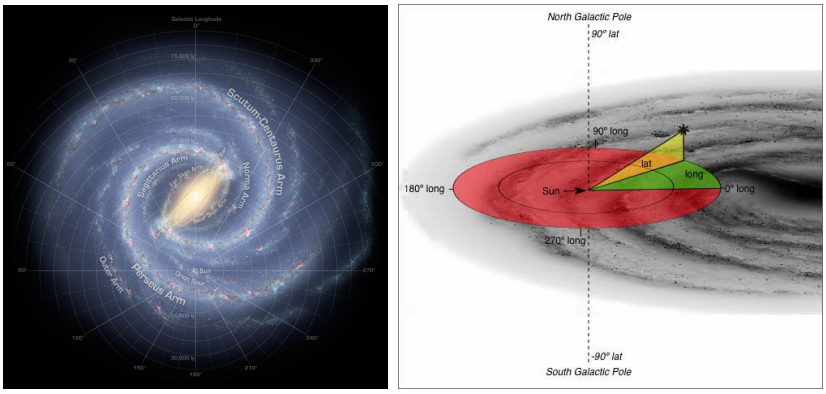


Figure 1.4: Galactic coordinates. The left panel shows an artist's rendition of what we think the Milky Way looks like. Galactic longitude is projected onto this image. The right panel shows a more edge-on view of Galactic longitude and latitude.

Component	Mass ($M_\odot$)
Total Mass	$\sim 10^{11}$
Stars	$\sim 5 \times 10^{10}$
Dark Matter	$\sim 5 \times 10^{10}$
Gas	$\sim 7 \times 10^{9}$
Dust	$\sim 7 \times 10^{7}$

Table 1.1: Mass of Milky Way components within 15 kpc of the Galactic center.

Most of the baryonic material in the Universe is actually ionized gas found at extreme temperatures in galaxy clusters. Weird, huh? In the Milky Way, this high-temperature plasma is much less important and we will not deal much with it.

The stellar and ISM distributions are flattened in the disk, but basically symmetric about the Galactic center (GC) and about the Galactic mid-plane. The MW disk is really thin, but has no sharp boundary in the z-direction - like our atmosphere on Earth. The gaseous component of the disk decreases to 50% of its mid-plane density at $\sim 250\,\mathrm{pc}$, so the full thickness is $\sim 500\,\mathrm{pc}$. This illustrates just how thin it really is. Another measure you may see is called the "scale height," which is the distance you have to travel for the density to decrease by a factor of e. This is obviously higher.

1.1.2 How do we know there is an ISM? (historical points shamelessly stolen from Harvard ay201b course notes)

Early astronomers pointed to 3 lines of evidence for the ISM:

- Extinction (absorption and scattering; also known as "attenuation"). Dust absorbs light from background stars. In 1919, Barnard called attention to these "dark markings" on the sky, and put forward the (correct) hypothesis that these were the silhouettes of dark clouds. A good rule of thumb for the amount of extinction present is 1

magnitude per kpc (for typical, mostly unobscured lines-of-sight; Wikipedia says 1.8; obviously varies quite a bit in the Galaxy). We will learn exactly what this means next class.

- Reddening (scattering; note that this is a part of extinction already). Even when the ISM doesn't completely block background starlight, it scatters it. Shorter-wavelength light is preferentially scattered, so stars behind obscuring material appear redder than normal. If a star's true color is known, its observed color can be used to infer the column density of the ISM between us and the star. Robert Trumpler first used measurements of the apparent "cuspiness" and the brighnesses of star clusters in 1930 to argue for the presence of this effect. Reddening of stars of "known" color is the basis of techniques used to map extinction today.

- In observations of binary stars whose spectral lines shift in velocity, astronomers noticed that some lines were not shifting. These lines were from stationary material between us and the binary system. Johannes Hartmann first noticed this in 1904 when investigating the spectrum of δ Orionis.

Helpful References: Good discussion of the history of extinction and reddening, from Michael Richmond (http://spiff.rit.edu/classes/phys230/lectures/ism_dust/ism_dust.html).

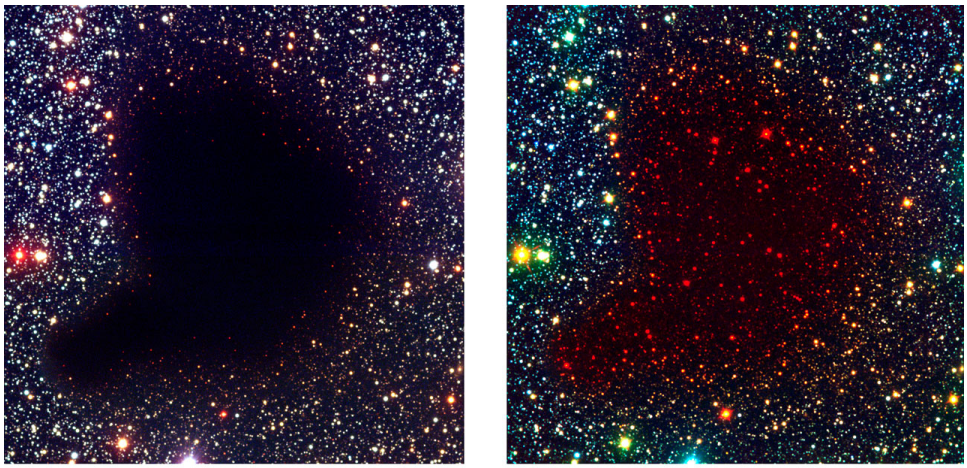


Figure 1.5: A small molecular cloud known as a "Bok globule." The left shows shorter wavelengths (B,V,I filters) than the right (B,I,K) filters. Notice how the longer wavelength image shows more stars (extinction), and those stars are redder (reddening).

1.1.3 Phases of the ISM (Draine Ch. 1.1)

We can break the ISM into various gas "phases," which each have characteristic densities and temperatures. These phases do not exist co-spatially, but can be assumed to be roughly in pressure balance with each other.

- Coronal gas - $\sim 10^5$ to 10^6 K plasma. Caused by SN. Also called "hot ionized medium"

(HIM). Dominates barionic matter out of the plane (and in galaxy clusters).

- H II gas - $\sim 10^4$ K plasma. Mostly caused by ultraviolet (UV) radiation from OB stars. H II regions such as Orion surround OB stars and live for MYrs. Planetary nebulae are created when evolved stars shove off their outer layers, and live for $\sim 10^4$ years. Diffuse $\sim 10^4$ K plasma is known as the "warm ionized medium" (WIM), and is maintained by OB stars but pervades the entire Galactic disk. Particle densities of $\sim 1\,\mathrm{cm}^{-3}$ for WIM and up to 10^6 for H II regions. [talk about cm^{-3}?]

- Warm H I- $T \sim 10^{3.7}$ K, density ~ 0.6 cm^{-3}. Also called "warm neutral medium" (WNM).

- Cool or Cold H I- $T \sim 10^2$ K, density ~ 30 cm^{-3}. Also called "cold neutral medium" (CNM).

- Diffuse molecular gas - Primarily H_2 but mainly traced by CO. Temperature of maybe 30 K (book says 50 K) and density of ~ 100 cm^{-3}.

- Dense molecular gas - Primarily H_2 but traced with many molecules (CS, NH_3, H_2CO, N_2H^+, etc.). Temperature of maybe 10 K (book says 10–50 K) and density of $\sim 10^3 - 10^6$ cm^{-3}. Makes stars.

Atoms and molecules of course change between these phases continually.

1.1.4 Elemental Composition (Draine Ch. 1.2)

Obviously, mostly H, but 40% He by mass (10% by number). After that, C, N, O, Ne, Mg, S are the main players, each with $\sim 10^{-5}$ the abundance of H, but together all "metals" with $Z > 3$ only add $\sim 1\%$ to the total mass. See Table 1.4. They nevertheless dominate the ISM temperatures due to cooling lines, and provide most of our diagnostics.

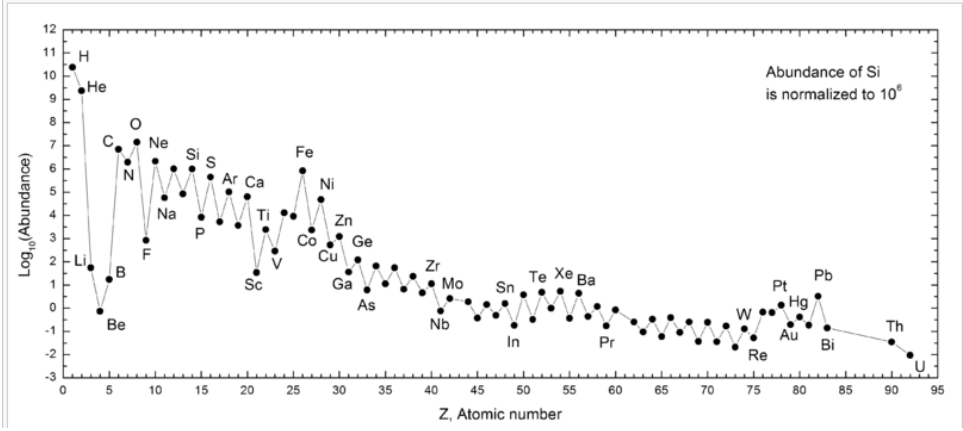


Figure 1.6: Elemental abundances.

1.1.5 Energy Densities

We can characterize the energy density in units of erg cm^{-3}. Many different forms:

- thermal ($u = 3/2nkT$)
- bulk kinetic or "ram pressure" ($u = 1/2\rho v^2$)
- CR (simply $u_{\rm CR}$ for now)
- magnetic ($B^2/8\pi$)
- EM (from CMB, dust, starlight, atoms, molecules, plasmas) $h\nu$ per photon, u_λ or u_ν for the radiation field, so $u = \lambda u_\lambda = \nu u_\nu$. For blackbodies, $u = aT^4$.

All of these energy densities are comparable! This is because the ISM self-regulates to some degree and energy is exchanged between the various sources. Too much of one would suppress another, leading to suppression of the former. Too much EM from stars would increase the thermal energy, making star formation more difficult, and decreasing EM. Too much CR would make CRs excape the Galaxy because the magnetic is too low, leading to a decrease in $u_{\rm CR}$.

[from http://www.ita.uni-heidelberg.de/ rowan/ISM_lectures/galactic-rad-fields.pdf] The EM field has some nomenclature that is important. Most of the EM field is from stars. Stars produce energy primarily in the near infrared, visible and soft ultraviolet. However, in neutral regions of the ISM, stellar photons with energies greater than the ionization energy of hydrogen, 13.6 eV, are largely absent - they are absorbed by hydrogen atoms, ionizing them, and hence cannot penetrate deeply into neutral regions.

We can characterize the EM field in various ways:

- Mathis et al. (1983, A&A, 128, 212) showed that in the solar neighborhood, the starlight component of the ISRF could be represented at long wavelengths as the sum of three diluted black-body spectra. At wavelengths $\lambda > 2450\,$Å, we have

$$\nu u_\nu = \sum_{i=1}^{3} \frac{8\pi h\nu^4}{c^3} \frac{W_i}{e^{h\nu/kT_i} - 1}\,, \tag{1.1}$$

  where W_i and T_i are the dilution factor and temperature of each component, with $T_1 = 3000$ K, $W_1 = 7.0 \times 10^{-13}$, $T_2 = 4000$ K, $W_2 = 1.65 \times 10^{-13}$, $T_3 = 7500$ K, $W_3 = 1.0 \times 10^{-14}$.

- At wavelengths $\lambda < 2450\,$Å, the starlight contribution to the ISRF has been estimated by a number of authors. The earliest widely-cited estimate was made by Habing (1968). He estimated that $\nu u_\nu \simeq 4 \times 10^{-14}$ erg cm^{-3} at $\lambda = 1000\,$Å, corresponding to a photon energy of 12.4 eV. It is often convenient to reference other estimates to this value, which

we do via the dimensionless parameter

$$\chi = \frac{(\nu u_\nu)_{1000\,\mathring{\mathrm{A}}}}{4 \times 10^{-14}\,\mathrm{erg\ cm^{-3}}} \tag{1.2}$$

- Alternatively, we can reference other estimates to the Habing field by comparing the total energy density in the range 6-13.6 eV. In this case, we define a different dimensionless parameter

$$G_0 = u(6 - 13.6\ \mathrm{eV}) 5.29 \times 10^{-14}\,\mathrm{erg\ cm^{-3}} \tag{1.3}$$

  If we are interested in e.g. the photodissociation of H_2 or CO, which requires photons with energies > 10 eV, then χ is the appropriate parameter to use. On the other hand, if we are interested in e.g. the photoelectric heating rate, which is sensitive to a wider range of photon energies, then G_0 is more appropriate.

- Two other estimates of the UV portion of the ISRF are in widespread use: one due to Draine (1978) and the other due to Mathis et al. (1983). Draine (1978) fit the field with a polynomial function:

$$\lambda u_\lambda = 6.84 \times 10^{-14} \lambda_3^{-5} (31.016\lambda_3^2 - 49.913\lambda_3 + 19.897)\,\mathrm{erg\,cm^{-3}}\,, \tag{1.4}$$

  where $\lambda_3 = \lambda/1000\,\mathring{A}$. This field has a normalization, relative to the Habing field, of $\chi = 1.71$ and $G_0 = 1.69$. Mathis et al. (1983) used instead a broken power-law fit:

$$\nu u_\nu = 2.373 \times 10^{-14} \lambda^{-0.6678} \quad 1340 - 2450\,\mathrm{\AA} \tag{1.5}$$

$$6.825 \times 10^{-13} \lambda \quad 1100 - 1340\,\mathrm{\AA} \tag{1.6}$$

$$1.287 \times 10^{-9} \lambda^{4.4172} \quad 912 - 1100\,\mathrm{\AA} \tag{1.7}$$

  Here, all wavelengths are in μm, and the energy densities are in units of erg cm^{-3}. This estimate has $\chi = 1.23$ and $G_0 = 1.14$. The available observational evidence suggests that the MMP83 estimate is a better one than that of Draine (1978), but the latter is probably in wider use in models of the ISM.

1.1.6 Putting it all together (see Wikipedia article on ISM as well)

Trends: T and n inversely correlated, and T and volume correlated (ideal gas law, yo!). T and scale height correlated.

1.2 Radiation (Kwok Ch. 2); see also ASTR601 notes

1.2.1 The Electromagnetic Spectrum

Astronomers have chopped up the EM spectrum and these terms get thrown around a lot - it is useful to review them here. These are all very approximate.

Compt.	Phase	Scale Height	Volume	Mass	Temp.	Density	How primarily observed
H I	CNM	100-300	$1-5$	CNM+WNM 3×10^9	100	100	21 cm H I line (absorption)
H I	WNM	300-400	$10-20$	CNM+WNM 3×10^9	$10^{3.7}$	0.6	21 cm H I line (emission)
H_2	Diffuse Molecular	100	<1	All H_2 1×10^9	~ 30	50-75	CO and OH radio lines, H_2 ro-vibrational IR lines, IR dust emission
H_2	Dense Molecular	50	<1	All H_2 1×10^9	~ 10	$\gtrsim 10^3$	Many lines in radio/IR, seen in emission and absorption, IR dust emission
H^+	HIM	1000-3000	$30-70$	HIM+WIM 1×10^9	10^6	$10^{-4}-10^{-2}$	X-rays (continuum and line), UV emission/absorption (lines)
H^+	WIM	50	$20-50$	HIM+WIM 1×10^9	10^4	10^2-10^4	Hα and optical/IR/radio lines, radio free-free
H^+	H II regions	50	$<1\%$	HIM+WIM 1×10^9	10^4	10^2-10^4	Hα and optical/IR/radio lines, radio free-free

Gamma Rays: above 100 keV per photon
"hard" X-rays: 10 to 100 keV per photon
"soft" X-rays: from $\sim$ 1 keV and 100 nm per photon
ultraviolet (UV): 100 - 300 nm per photon
"far" UV (FUV): 100-200 nm per photon
"near" UV (FUV): 200-300 nm per photon
optical (visible): 300-700 nm (UVBRI filters; well maybe not U or I)
Infrared (IR): 700 nm to 500 μm
Near IR (NIR): 500 nm to 3 μm(JHK filters)
Mid IR (MIR): 3 to 50 μm
Far IR (FIR): 50 to 500 μm
Sub-mm: 500 μm to 1 mm
mm: 1-3 mm
Radio: below 3 mm.

Each of these regimes has "bandpasses" or filters that define wavelengths where observations are conducted. Here are a few:

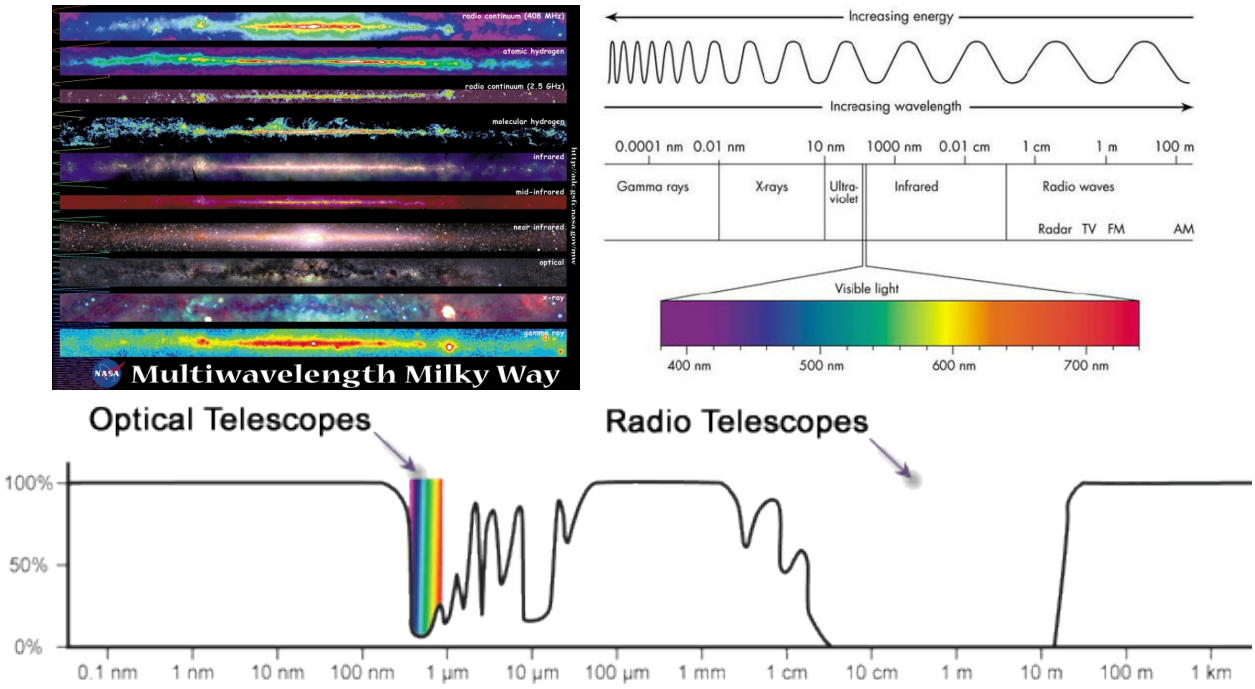


Figure 1.7: The Milky Way midplane seen at a variety of wavelengths (top left), the EM spectrum (top right), and the opacity curve of the Earth's atmosphere (bottom; 0% is opaque and 100% is complete transmission).

Sometimes these bandpasses correspond to windows in the atmostphere (see ALMA bandpasses), and sometimes they are determined by the materials used to create the filter (the HST ones).

Most observations can be categorized as either "photometry" and "spectroscopy." In pho-

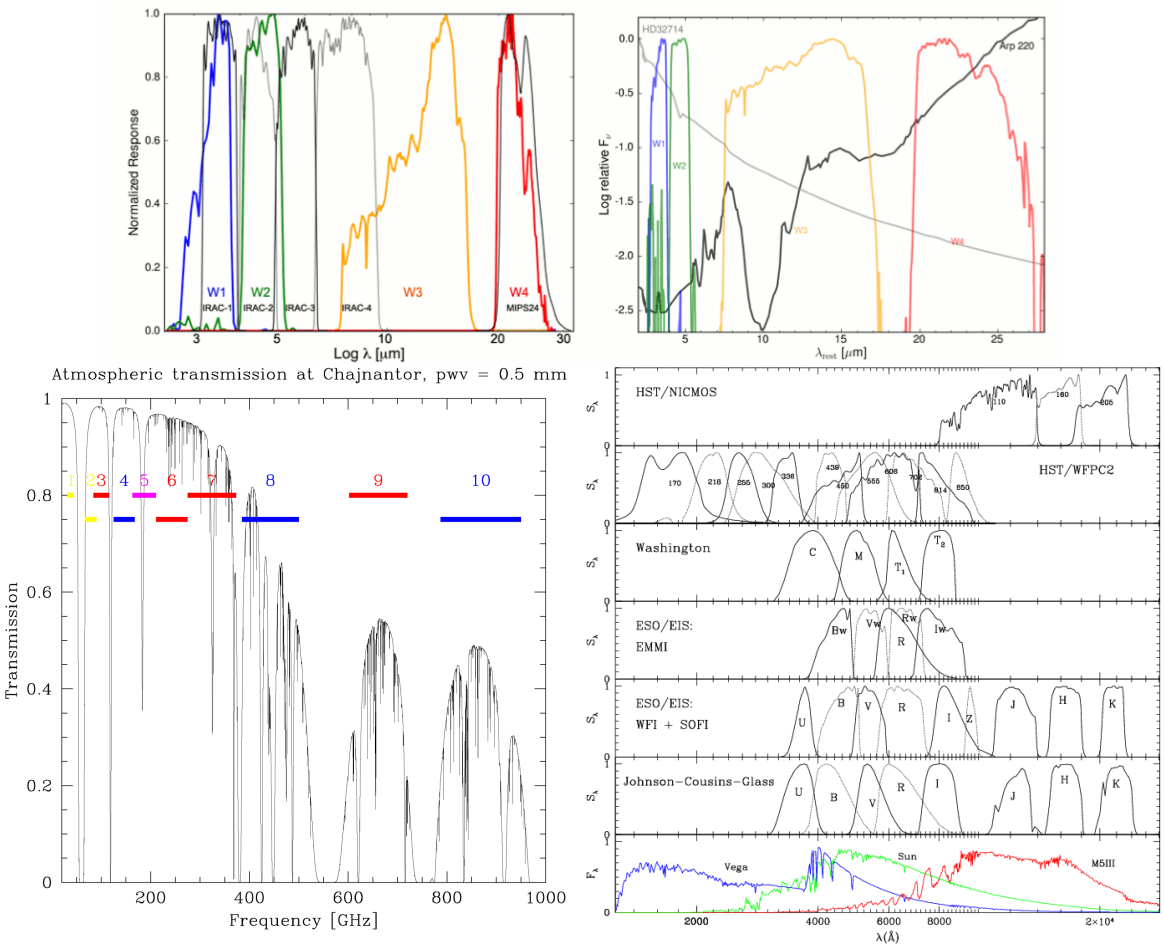


Figure 1.8: Filter bandpasses for infrared observatories *Spitzer* and *WISE* (top), ALMA (bottom left) and various optical observatories (bottom right).

tometry, you get an image of the flux, averaged over a bandpass. In spectroscopy, you get a spectrum, or a 1D range of values as a function of wavelength. You can also do photometry of spectral lines, which is called "narrow-band" photometry.

1.2.2 Intensity

The *specific intensity* of radiation the most basic observable quantity. It is essentially the surface brightness, and is appropriate for all resolved objects.

Two important things about brightness (from NRAO radio course):
(1) As long as the source is resolved, intensity is independent of distance. Thus, the camera exposure time and aperture setting for an exposure of the Sun would be the same, regardless of whether the photograph was taken close to the Sun (from near Venus, for example) or far away from the Sun (from near Mars, for example), *as long as the Sun is resolved in the photograph*. This seems terribly wrong at first, but can easily be proven.

(2) Intensity is the same at the source and at the detector. Thus you can think of brightness in terms of energy flowing out of the source or as energy flowing into the detector.

In radio astronomy, we use units for specific intensity of temperature in Kelvin, or in units of $\mathrm{Jy\,beam^{-1}}$. In the MIR and FIR, a common units for intensity is $\mathrm{MJy\,sr^{-1}}$. At visible wavelengths, you may simply see photon counts.

Intensity is related to the energy dE passing through an infinitesimally small area $d\sigma$ by:

$$dE = I_\nu \, d\sigma \, \cos\theta \, d\Omega \, d\nu \, dt \, . \tag{1.8}$$

Here, "specific" refers to the fact that it is at a particular wavelength. We can of course rewrite this as:

$$I_\nu = \frac{dE}{d\sigma \, \cos\theta d\Omega d\nu dt} \, . \tag{1.9}$$

θ is measured normal to the surface $d\sigma$ and $d\Omega$ is the solid angle. The dimensions of I_ν are then $\mathrm{erg\,cm^{-2}\,Hz^{-1}\,s^{-1}\,sr^{-1}}$.

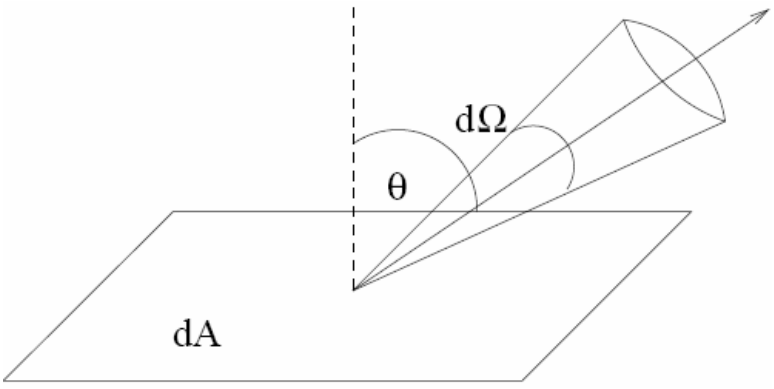


Figure 1.9: Geometry of an astronomical observation. $d\Omega$ corresponds to the solid angle and dA is the area of the detector.

Notice that we wrote the specific intensity in frequency units. I_ν has a dependence on $d\nu$, and $d\nu \neq d\lambda$. Instead,

$$d\nu = -(c/\lambda^2) d\lambda \tag{1.10}$$

so combining with the above equations

$$\nu I_\nu = \lambda I_\lambda \tag{1.11}$$

To get the *intensity* or *integrated intensity* we would of course integrate over frequency or wavelength:

$$I = \int_0^\infty I_\nu d\nu = \int_0^\infty I_\lambda d\lambda \tag{1.12}$$

Because integrating over frequency always requires many assumptions, these quantities are rarely used.

1.2.3 Solid Angles

The discussion of specific intensity above included solid angles, which many students haven't yet heard of. A solid angle, measured in dimensionless steradians (sr), is simply a two-dimensional angle. Think of it as a cone spreading out from the center of a sphere to its edge. A solid angle is the area of a unit sphere such that there are 4π sr total on a sphere. The obvious application is the sky. Objects that appear larger on the sky have a larger solid angle.

The mathematical definition is

$$d\Omega = \sin\theta d\theta d\phi \tag{1.13}$$

or

$$\Omega = \int_S \int \sin\theta d\theta d\phi\,, \tag{1.14}$$

where θ and ϕ are angles in spherical coordinates and the integration is over surface S.

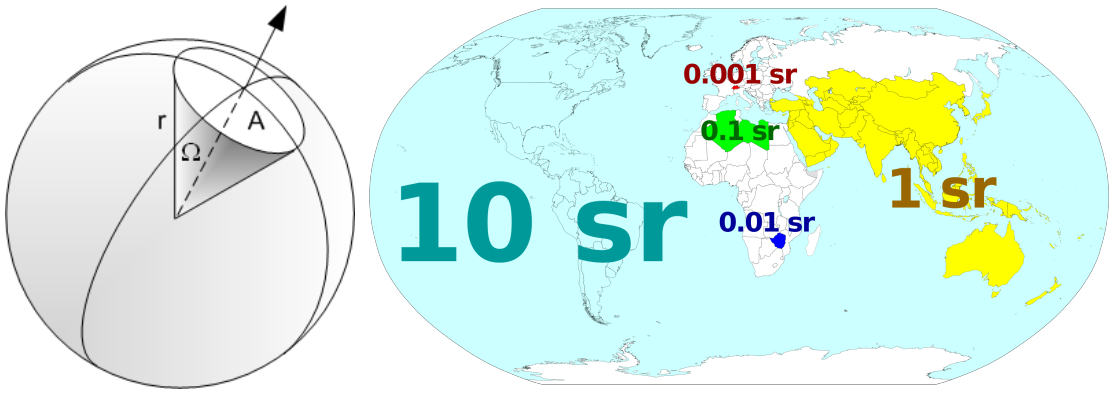


Figure 1.10: Solid angles. The left image shows the relationship between the area on the sphere and the solid angle Ω. The right panel shows some values of the solid angle of the Earth. The entire Earth is 4π sr, or 12.57 sr.

For a spherical solid angle, $\theta = \phi$. For spherical solid angles with small θ we can approximate the solid angle with:

$$\Omega \simeq \pi\theta^2\,, \tag{1.15}$$

with θ in radians of course. Notice that this is just the area of a circle of radius θ. The true solid angle will be slightly smaller than this for a given value of θ, although this is almost always appropriate for astronomical measurements. The true formula is

$$\Omega = 2\pi(1 - \cos\theta) \tag{1.16}$$

Problem: What is the solid angle of a single $6''$ square pixel?

1.2.4 Flux

While intensity is perfect for extended sources, we are frequently more interested in the quantity of *flux* integrated over solid angle:

$$F_\nu = \int^{d\Omega} I_\nu \cos\theta d\omega \tag{1.17}$$

or

$$F_\nu = \int_0^{2\pi}\int_0^{\pi} I_\nu \cos\theta \sin\theta d\theta d\phi\,. \tag{1.18}$$

The units of flux are therefore $\mathrm{erg\,cm^{-2}\,s^{-1}\,Hz^{-1}}$. Note that this is technically the "specific" flux, but we don't use that term for some reason - instead we call this term the "flux density." This is confusing terminology, I know! In the radio and MIR/FIR regimes, we frequently use the unit of Jansky (Jy), which is $10^{-23}\,\mathrm{erg\,cm^{-2}\,s^{-1}\,Hz^{-1}}$. Similar to the intensity, we can integrate over frequency or wavelength to (again) get the *flux* or *integrated flux*. Again, this is rarely done.

In practice, when do we use spectral brightness and when do we use flux density to describe a source? If a source is unresolved, meaning that it is much smaller in angular size than the point-source response of the eye or telescope observing it, its flux density can be measured but its spectral brightness cannot. If a source is much larger than the point-source response, its spectral brightness at any position on the source can be measured directly, but its flux density must be calculated by integrating the observed spectral brightnesses over the source solid angle.

1.2.5 Luminosity

Intensity and flux are observable quantities, and not physical quantities. We are often more interested in luminosity, which is intrinsic to the source. The *observed* luminosity is

$$L_\nu = 4\pi d^2 F_\nu\,, \tag{1.19}$$

where d is the distance to the source. Again, even if it's a "specific" luminosity, we generally just say "luminosity."

Luminosity can also be defined leaving a surface:

$$dL_\nu = \int_{\mathrm{surface}} F_\nu dA\,. \tag{1.20}$$

Since for black bodies, $F = \sigma T^4$ (Stephan-Boltzmann), $L = A\sigma T^4$ if the object can be characterized by a single temperature. [This information repeated below in the blackbodies section.]

Notice that I have given you two different definitions of luminosity! One is appropriate when you know the emergent flux from a surface, and one is appropriate when you know the observed flux that you measure with your telescope, similar to how we discussed intensity. Both are fluxes and have the same units. Because both flux definitions almost never show up together in an expression, I have neglected to differentiate between them.

1.2.6 Magnitudes

Magnitudes are the units of brightness, typically used in the optical and near-infrared regimes. They are always measured in a particular bandpass, for example the Johnson V-band (see bandpasses earlier in this lecture). This allows us to compute "colors" by looking at magnitude differences. Colors are the crudest way of determining the shape of the spectral energy distribution.

The study of the ISM rarely requires use of magnitudes, since magnitudes are most often used to characterize stars. Stars are not part of the ISM! I'll include information on magnitudes here for completeness.

Magnitudes are based on Hipparchus's classification of stars in the northern sky. Hipparchus classified stars with values of magnitudes from 1 to 6, 1 magnitude being the brightest. Because it was defined by eye, and the eye does not have a linear response, a first magnitude star is not twice as bright as a second magnitude star. Instead, astronomers later found that Hipparchus' system is roughly logarithmic, and 6th magnitude stars are roughly 100 times fainter than 1st magnitude stars. The magnitude system has two peculiarities:
(1) It is defined backwards, and
(2) It is logarithmic.
So it is basically the ideal system to use.

Five equal steps in log-space (1st to 6th magnitude) result in factors of 2.512 in linear space ($100^{\Delta m/5} = 2.512^{\Delta m}$). Therefore, a 1st magnitude star is 2.512 times brighter than a second magnitude star, and a 4th magnitude star is $2.512^3 = 15.8$ times fainter than a 1st magnitude star. Another way of thinking about this is:

$$m_1 - m_0 = -2.5 \log_{10}(F_1/F_0) \tag{1.21}$$

or

$$\frac{F_1}{F_0} = 10^{0.4(m_0 - m_1)}\,, \tag{1.22}$$

where F_1 and F_0 are the fluxes and m_0 and m_1 are the magnitudes at wavelengths or frequencies "0" or "1".

How can we actually use this system? We need a reference star of known flux and magnitude. Any star will do, but two commonly used ones are Vega and the Sun.

Instead of arbitrary wavelengths, we usually use photometric filters. The most common filters used are the Johnson *U,V,V,R,I*, but there are now a large number of filters available. Magnitudes found using these filters are often denoted with the filter names themselves, e.g., B for m_B.

Question: a star has a $B-V$ color of -1. Stars with $B-V=0$ appear slightly blue. Does this star look more or less blue? What does that imply about its temperature?

In the ISM, dust attenuates star light, and this attenuation is often measured in magnitudes. Each kpc in the Galaxy produces about a magnitude of visual extinction. Star formation

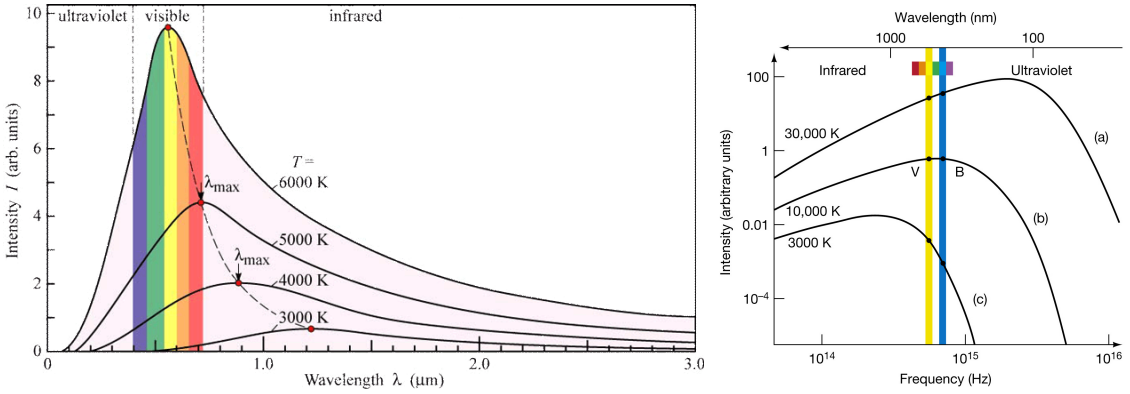


Figure 1.11: Blackbody curves in linear (left) and log (right) space.

regions can have visual extinctions of 100, so a star would have $2.512^{100} = 10^{40}$ times less light than it would if extinction were not present. Extinction generally decreases with increasing wavelength, so it is less in the infrared and essentially absent in the radio. We will discuss the exact form of the extinction law in some detail.

1.2.7 Blackbody Emission

You all know blackbodies (Planck functions):

$$B_\nu = \frac{2h\nu^3}{c^2}\frac{1}{e^{h\nu/kT}-1}\,. \tag{1.23}$$

or

$$B_\lambda = \frac{2hc^2}{\lambda^5}\frac{1}{e^{hc/\lambda kT}-1}\,. \tag{1.24}$$

But what you may not know are the units: erg cm^{-2} s^{-1} Hz^{-1} sr^{-1} for B_ν or erg cm^{-2} s^{-1} cm^{-1} sr^{-1} for B_λ, where the additional "cm" term is the wavelength (often given in Angstroms). In other words, it is a specific intensity! But under what conditions is $I_\nu = B_\nu$? We'll see in a bit.

If we are on the right side of the peak in the long wavelength limit, $e^{hc/\lambda kT} - 1 \simeq 1 + hc/\lambda kT - 1 = hc/\lambda kT$. In frequency units, we find $e^{h\nu/kT} - 1 \simeq 1 + h\nu/kT - 1 = h\nu/kT$. We can therefore write

$$B_\lambda = \frac{2ckT}{\lambda^4} \tag{1.25}$$

or

$$B_\nu = \frac{2\nu^2 kT}{c^2} \tag{1.26}$$

This is known as the *Rayleigh-Jeans limit* or *Rayleigh-Jeans approximation*. [When is this approximation valid?] We almost always assume this limit in radio astronomy.

$$S_\nu = \int_{beam} I_\nu d\Omega\,, \tag{1.27}$$

where S_ν is the flux density [radio astronomy doesn't use F_ν for some reason], and the integration is over the telescope beam (not necessarily over the entire source!). We can refer to the temperature as the *brightness temperature*, T_B, and

$$S_\nu = \int_{beam} B_\nu d\Omega = \int_{beam} \frac{2\nu^2}{c^2} k T_B d\Omega \,. \tag{1.28}$$

What is the brightness temperature? It is the value that is needed to give the measured flux S_ν. Wikipedia's definition: "Brightness temperature is the temperature a black body in thermal equilibrium with its surroundings would have to be to duplicate the observed intensity of a grey body object at a frequency." This is of course not necessarily the kinetic temperature. If the source of interest is a blackbody this is the kinetic temperature, $T_B = T_k$, but this is not necessarily true in all cases. *If* the source has a constant surface brightness over the telescope beam,

$$S_\nu = \frac{2\nu^2}{c^2} k T_B \Omega \,. \tag{1.29}$$

There are a two important points about blackbody radiation. First, we can use Wien's Law to determine the peak wavelength (or frequency):

$$\lambda_{\rm max} = \frac{0.2898}{T(\,{\rm K})} \ {\rm cm}\,, \tag{1.30}$$

or

$$\nu_{\rm max} = 5.879 \times 10^{10} T(\,{\rm K})\,. \tag{1.31}$$

We can derive these by setting the differential of B_λ or B_ν equal to zero. This tells us that hotter things peak at shorter wavelengths and higher frequencies. This illustrates the utility of colors, but only if you roughly know the temperature of the object!

In the infrared, we have some handy rules of thumb: a 30 K cloud peaks at 100 μm, and a 100 K cloud will peak at 30 μm. Hot stars (30000 K) peak at 100 nm in the UV. The Sun (6000 K) peaks at 500 nm in the visible (green).

A second important point is that a hotter blackbody emits more intensity at all wavelengths. This can be see in the above figures. Since the blackbody units are those of intensity, this means that a hotter blackbody has a higher surface brightness than a cooler blackbody.

It is important to remember that more intensity at all frequencies does not necessarily mean more energy! Think about burners on a stove. A small hot burner will have very intense radiation. A large cooler burner will have less intense radiation. But the larger one may boil water faster because although its intensity (surface brightness) is lower, it emits more total energy.

Let's quantify this. To find the intensity (not the specific intensity), we integrate over all frequencies or wavelengths:

$$B(T) = \int_0^\infty B_\nu(T) d\nu \,. \tag{1.32}$$

After some math, this integral results in the expression

$$B(T) = \frac{\sigma T^4}{\pi}\,, \tag{1.33}$$

where σ is of course the Stephan-Boltzmann constant. In the case of an isotropic radiation field, which we can almost always assume, it can be shown that $F = \pi I$, so therefore $F = \sigma T^4$. This is of course the *Stephan-Boltzmann Law* again. We are often interested in the total luminosity of an object (in erg s^{-1} or W):

$$L = \int_S F dA\,, \tag{1.34}$$

the flux integrated over the emitting surface. For spherical objects, this leads to $L = 4\pi r^2 \sigma T^4$, where r is the object's radius. Thus, the total energy output is related to the surface area and the temperature.

1.2.8 Types of Emission

There are broadly two types of emission: continuous and line. For continuous emission, there are no quantized energy states, and the particles can therefore emit at any frequency. "Line" radiation, either in emission or absorption, arises from transitions between discrete energy states.

Examples of continuous emission are: blackbodies, free-bound, bound-free, and Brehmsstrahlung. Examples of line emission are H I 21 cm, CO rotational transitions, etc. These are not mutually exclusive! For example, plasmas emit continuous Brehmsstrahlung emission, as well as recombination line emission.

Line emission requires an understanding of atomic energy states. To express these states, we use spectroscopic notation:

$$^{2S+1}L_J\,, \tag{1.35}$$

where

- $\vec{S}$ is the total spin quantum number and $2\vec{S}+1$ is the number of spin states; $\vec{S}$ is the absolute value of the total electron spin $\vec{S} = |(\sum s_i)|$; $\vec{S} = 0$ for a closed shell and $\vec{S} = 1/2$ for hydrogen;
- $\vec{L}$ is the total orbital angular momentum ($\vec{L} = \sum \vec{\ell_i}$), which is written as $S, P, D, F, ...$ for $|\vec{L}| = 0, 1, 2, 3...$
- $\vec{J}$ is the total angular momentum quantum number; $\vec{J} = \vec{L} + \vec{S}$; for a given $\vec{L}$ and $\vec{S}$, there are $(2L+1)(2S+1)$ possible values of J. For a hydrogenic ion, $L = 0$, $S = 1/2$, and $J = 1/2$. For more complex atoms, J takes on the values $L+S$, $L+S-1$, ..., $|(L-S)|$.

The splitting of energies based on total angular momentum $\vec{J}$ is known as "fine-structure" splitting. Transitions between these different energy states defined by J-levels are not allowed by electric dipole radiation, but are very important in the ISM. There are further splittings called "hyperfine structure," based on the interaction between electronic and nuclear spins. This is not captured in spectroscopic notation.

And two new vocabulary words: The "term" is the combination characterized by a specific S and L; The "level" is the combination characterized by specific values of L, S, and J. The difference in the energy between two levels gives the wavelength or frequency of an atomic transition.

The lowest state of hydrogen is $S = 1/2, L = 0, J = 1/2$, or ${}^2S_{1/2}$. The ground state of Boron has a ${}^2P_{1/2}$ term. Closed shells always have a ${}^1S_0$ term. As an example, for $S = 1, L = 2$, there are $(2 \times 1 + 1)(2 \times 2 + 1) = 15$ different levels corresponding to the ${}^3D$ term, of which $(2 \times 3 + 1) = 7$ belong to the ${}^3D_3$ ($J = 3$) level.

1.2.9 The Interstellar Radiation Field

Because dust emits and absorbs (largely) continuous radiation, it interacts much more strongly with the interstellar radiation field compared to gas. At this point it's good to review our notes on the interstellar radiation radiation field (ISRF). These are contained in Draine Chapter 12, and in the introduction notes. Below, we'll expand on those notes.

Briefly, the ISRF has six components:
1) Synchrotron radiation from relativistic electrons
2) CMBR
3) IR emission from dust grains
4) Thermal emission from warm (10^4) plasma
5) Starlight
6) X-ray emission from hot (10^5 to 10^8 K plasma)

Draine Chapter 12 goes into some detail on these six components, in order of increasing frequency.

Galactic Synchrotron Radiation

We will discuss synchrotron in more detail when we talk about plasmas. Synchrotron radiation is known as "non-thermal," which means that the intensity of the radiation does not depend on the temperature of the source. Synchrotron dominates at the lowest radio frequencies, and its spectrum can be defined as

$$\nu u_\nu \approx 2.86 \times 10^{-19} \nu_9^{0.05} \,\mathrm{erg\,cm^{-3}}\,, \tag{1.36}$$

between 400 MHz and 1.4 Hz for $|b| > 5$ deg, where ν_9 is the frequency in GHz. The synchrotron intensity is the same as that of the CMBR near ~ 1 GHz. AGN and supernova remnants emit synchrotron radiation.

CMBR

The CMB is essentially isotropic radiation from the Big Bang, with a temperature of $2.7255 \pm 0.0006\,\mathrm{K}$. The CMB is slightly anisotropic though, due to the Sun's motion relative to the CMB rest frame. Dominant source of intensity from $\sim 1\,\mathrm{GHz}$ to a few mm.

Free-Free and Recombination Continuum

Warm ($\sim 10^4\,\mathrm{K}$) plasma has various emission mechanisms, which are mainly from free-free and free-bound. Both of these are "continuous," because they involve un-quantized energy states. We will go into these emission mechanisms in much more detail later. These mechanisms are not usually dominant at any wavelength.

Dust

Dust dominates the emission spectrum from $\sim 600\,\mu\mathrm{m}$ to $5\,\mu\mathrm{m}$. About 2/3 of the power is from wavelengths greater than $50\,\mu\mathrm{m}$. This is the "gray-body" emission from large grains. Gray-bodies are just blackbodies with slightly lower optical depths. The peak of this radiation near $1\,\mu\mathrm{m}$ tells us that the dusts is near $30\,\mathrm{K}$. Dust dominates from a few mm to a few μm.

Dust also has pseudo line emission seen in Draine Figure 12.1. These arise from bending and stretching modes of smaller grains, most prominently polycyclic aromatic hydrocarbons. Their emission dominates from ~ 3 to $\sim 20\,\mu\mathrm{m}$.

Starlight

Starlight can take the form of a blackbody. As you hopefully found in your first homework, the combined spectrum from a population of stars peaks near $1\,\mu\mathrm{m}$. This is the emission from a star with temperature near $3000\,\mathrm{K}$.

There are a few important points about starlight:
Larger stars emit more UV light (hotter Blackbodies). Therefore, the local ISRF changes depending on location.
Photons with $E > 13.6\,\mathrm{eV}$ are quickly absorbed by the ever-present hydrogen.

FUV radiation (1) photo-excites and photo-dissociates H_2 and other molecules, (2) photo-ionizes heavy elements, and (3) ejects photo-electrons from dust grains.

We can classify UV radiation fields in terms of Habing units (Habing, 1968):

$$\chi = \frac{(\nu u_\nu)_{1000}}{4 \times 10^{-14}\,\mathrm{erg\,cm^{-3}}}\,. \tag{1.37}$$

the 4×10^{-14} number is from Habing (1986), and χ is known as "Habings." The intensity is measured near 1000 Angstroms in the UV.

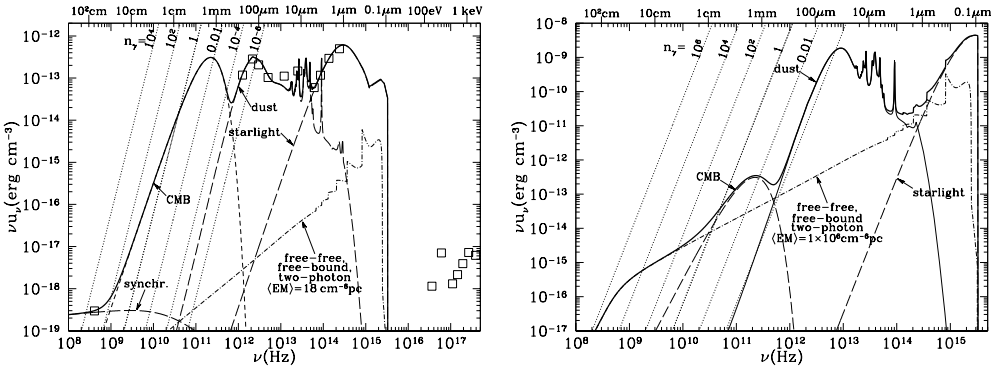


Figure 1.12: (left) ISRF near the Sun (Draine 12.1). Spectral lines not included. Squares show measurements. Dotted lines show constant photon occupation number n_γ. (right) Same as Figure 5.24 (Draine 12.3), but near an O star with $T = 3.5 \times 10^4$ K. Again, spectral lines not shown and dotted lines are constant n_γ.

Sometimes a wider range of UV radiation is better, so we also have the parameter G_0, which is referenced to the Habing (1968) spectrum integrated between 6.0 and 13.6 eV, which yields $5.29 \times 10^{-14}\,\mathrm{erg\,cm^{-3}}$. G_0 is therefore:

$$G_0 = \frac{u(6 - 13.6)\,\mathrm{eV}}{5.29 \times 10^{-14}}\,\mathrm{erg\,cm^{-3}} \tag{1.38}$$

You will see values quoted for Habing units and G_0.

X-Rays from Hot Plasma

Just as we get radio emission from warm plasma, we get X-Ray emission from hot (10^6 K) plasma. Where does this hot plasma come from? In our Galaxy, it is mainly from supernovae. In large galaxy clusters, it is from electrons "falling" in to the potential and accelerating.

In our Galaxy, X-rays play a very small role. In a cluster of galaxies, they dominate the emission.

Note that

$$n_\gamma = \frac{c^2}{2h\nu^3} I_\nu \tag{1.39}$$

or

$$\overline{n_\gamma} = \frac{c^3}{8\pi h\nu^3} u_\nu \tag{1.40}$$

for the average over all directions (factor of 4π).

Views of the sky: http://mwmw.gsfc.nasa.gov/mmw_allsky.html

1.3 Radiative Transfer [see ASTR601 notes]

1.3.1 The Equation of Radiative Transfer

Radiative transfer is the change in intensity dI_ν as radiation propagates from a source (I_ν) to the observer. [Note: we will use a "0" subscript in future subsections to denote the background but will drop it in this subsection for clarity.] Along the way, the emission will either be absorbed and scattered by intervening material, or it will encounter an emitting region.

For attenuation, we can define a "linear absorption coefficient" κ_ν with units of cm^{-1}. *This is misleading since it contains contributions from both absorption and scattering!* Note that this is not opacity or mass absorption coefficient, although both share the same notation and a similar definition! Sorry for the confusion. The amount of energy absorbed is proportional to the light intensity:

$$dI_\nu = -\kappa_\nu I_\nu ds\,, \tag{1.41}$$

where ds is the path. Absorption removes photons from the path, thus the negative sign. It is worth pointing out here that absorption excites atoms and molecules, and these atoms and molecules then re-emit. If this emission were beamed along ds there would be no change in intensity. Instead, the re-emitted light is more generally close to isotropic, so the emission is reduced.

For emission, we can define the emission coefficient (or *emissivity* j_ν as:

$$dI_\nu = j_\nu ds\,. \tag{1.42}$$

Notice that there is no dependence on $I_{\nu,0}$, in contrast to absorption. The units of j_ν are $\mathrm{erg\,cm^{-1}\,sr^{-1}\,s^{-1}}$.

The total change in intensity is therefore

$$dI_\nu = j_\nu ds - \kappa_\nu I_\nu ds\,, \tag{1.43}$$

or

$$\frac{dI_\nu}{ds} = j_\nu - \kappa_\nu I_\nu\,. \tag{1.44}$$

This is one form of the *Equation of Radiative Transfer*. This is one of the fundamental equations in astrophysics. All it is saying, however, is that the change in intensity along the path is just the emission (j_ν) minus the absorption ($\kappa_\nu I_\nu$).

Let's take the illustrative example of no emission. In this case

$$\frac{dI_\nu}{ds} = -\kappa_\nu I_\nu\,, \tag{1.45}$$

which has a solution

$$I_\nu(s) = I_{\nu,0} e^{-\kappa_\nu s}\,, \tag{1.46}$$

where $I_{\nu,0}$ is the unattenuated emission. The radiation intensity will decrease exponentially.

We can also define the dimensionless quantity of *optical depth* τ_ν from

$$d\tau_\nu = \kappa_\nu \, ds \, . \tag{1.47}$$

or

$$\tau_\nu = \int \kappa_\nu \, ds \, , \tag{1.48}$$

where the integration is carried out over the path length. In most cases, we need only integrate over the source of interest. For example, if there is a gas cloud 20 kpc away that is 1 kpc thick, we may be able to only integrate over the 1 kpc of the cloud if the rest of the 20 kpc can be assumed to have no impact. For completeness,

$$I_\nu(\tau_\nu) = I_{\nu,0} e^{-\tau_\nu} \, , \tag{1.49}$$

The optical depth ranges from zero to infinity.
Low values $\tau_\nu \ll 1$ are called "optically thin." These are things you can see through *at that particular frequency*. A good example is glass, which has a very low optical depth at optical frequencies, but actually has a high optical depth in the ultra-violet.
High values $\tau_\nu \gg> 1$ are called *optically thick*. A wall is optically thick at optical frequencies. A wall is optically thin at X-ray frequencies.
Near $\tau \simeq 1$ we have to be careful - this is marginally optically thick.

If we rewrite things in terms of the optical depth, using $\frac{d\tau_\nu}{ds} = \kappa_\nu$,

$$\frac{dI_\nu}{d\tau_\nu} = \frac{j_\nu}{\kappa_\nu} - I_\nu \, . \tag{1.50}$$

We can further define the *Source function* S_ν

$$S_\nu = \frac{j_\nu}{\kappa_\nu} \tag{1.51}$$

Combining our expressions, we arrive a second form of the *Equation of Radiative Transfer*, this time using the optical depth and source function:

$$\frac{dI_\nu}{d\tau_\nu} = S_\nu - I_\nu \tag{1.52}$$

We will use this one from now on, because optical depth is a much better and more measurable parameter compared with actual linear depth.

In (full) thermodynamic equilibrium (TE) at temperature T, there is no change in intensity along the path and $\frac{dI_\nu}{d\tau_\nu} = 0$. In this case, $I_\nu = S_\nu = B_\nu(T)$, our old friend the Planck function. When is $I_\nu = B_\nu(T)$??? When $d\tau \to \infty$! Or in other words, when the optical

depth is high, the intensity is that of a blackbody at temperature T. In this case, nothing else about the source matters, only its temperature.

This is a subtle, but extremely important point. For high optical depth sources, the only emission you can get out is that of a blackbody. You cannot for example get line emission. The source properties, aside from temperature, do not matter. The only thing you see is the *surface* emission. In fact, you only see down on average to the depth where the optical depth is unity. Think of a wall again, where you cannot determine how thick it is since you only see the paint layer (ok, so a wall actually is not a perfect blackbody since paint reflects light of different wavelengths....). Contrast this with glass. As glass get thicker, and thicker, we will notice more of a green hue. By determining *how* green it is, we can work out how thick it is. We will return to this point later.

1.3.2 Solutions to the Equation of Radiative Transfer

The deceptively simple equation of radiative transfer has had volumes written about its solutions. We can integrate the transfer function by multiplying by e^{τ_ν}:

$$e^{\tau_\nu}(dI_\nu + I_\nu d\tau_\nu) = e^{\tau_\nu} S_\nu d\tau_\nu \tag{1.53}$$

$$d(e^{\tau_\nu} I_\nu) = e^{\tau_\nu} S_\nu d\tau_\nu \tag{1.54}$$

If we define $\tau_\nu = 0$ at $I_{\nu,0}$,

$$e^{\tau_\nu} I_\nu - I_{\nu,0} = \int_0^{\tau_\nu} e^{\tau'} S_\nu d\tau' \tag{1.55}$$

multiply by $e^{-\tau_\nu}$

$$I_\nu(\tau_\nu) = I_{\nu,0} e^{-\tau_\nu} + \int_0^{\tau_\nu} S_\nu(\tau') e^{-(\tau_\nu - \tau'_\nu)}\, d\tau' \tag{1.56}$$

The intensity I_ν at optical depth τ_ν is the initial (background) intensity $I_{\nu,0}$ attenuated by a factor $e^{-\tau_\nu}$, plus the emission $S_\nu d\tau'$ integrated over the path, itself attenuated by the factor $e^{\tau_\nu - \tau'_\nu}$. This final exponent represents "self-absorption." The material itself will absorb its own radiation. "Self-absorption" refers to absorption by one species (H I, CO, etc) by that species. If background radiation from e.g. H I is absorbed by optically thick H I, this is called self-absorption. We will revisit this when we talk about H I and CO. This is known as the "formal solution to the equation of radiative transfer."

The difficulty in using Equation 1.56 is that in general we don't know how S varies with τ, because S depends on I, which is not known until S is known. It's a circular problem, which is why it is often solved computationally. It is worth examining this equation a bit more in limiting cases that allow us to simplify the integral:

$\tau = 0$

If the optical depth is zero, we get $I_\nu = I_{\nu,0}$, simply the background intensity back. If there is no optical depth, we get neither emission nor absorption (like a window!). This illustrates how emission and absorption are intimately related.

$S = 0$

If the source function is zero, we get $I_\nu = I_{\nu,0} e^{-\tau_\nu}$, which is frequently written as $I_\nu / I_{\nu,0} = e^{-\tau_\nu}$. This is an absorption line! As τ_ν increases, the absorption becomes stronger.

S constant

We can sometimes make the assumption that S_ν is a constant, so we can pull it out of the integral. Thus:

$$I_\nu(\tau_\nu) = I_{\nu,0}(\tau_\nu) e^{-\tau_\nu} + S_\nu \int_0^{\tau_\nu} e^{(-\tau - \tau')} \, d\tau' = I_{\nu,0}(\tau_\nu) e^{-\tau_\nu} + S_\nu (1 - e^{-\tau_\nu}) \tag{1.57}$$

This is the most useful general form of the solution of radiative transfer for our purposes. The first term on the right hand side is attenuation by the ISM along the line of sight. The second one is emission from the ISM along the line of sight.

S constant, LTE

In Local Thermodynamic Equilibrium, LTE, $S_\nu = B_\nu(T)$ (we'll revisit this), so

$$I_\nu(\tau_\nu) = I_{\nu,0} e^{-\tau_\nu} + B_\nu (1 - e^{-\tau_\nu}) \tag{1.58}$$

We will discuss LTE later, but essentially it means that for a small volume we can assume a single temperature that is also reflected in the level populations of the atoms and molecules.

S constant, LTE, Radio Regime

In the radio, we use the brightness temperature instead of the intensity. They are related by $I_\nu = \frac{2\nu^2}{c^2} k T_B$, where T_B is the "brightness temperature," essentially the intensity the ideal radio telescope would measure (we frequently drop the "B" subscripts in this expression). We can also use the Rayleigh-Jeans approximation $B_\nu(T) = \frac{2\nu^2}{c^2} kT$, with T here the kinetic temperature. Since these relationships both have the same constants, we can write

$$T_B = T_{B,0} e^{-\tau_\nu} + T(1 - e^{-\tau_\nu}) \tag{1.59}$$

Note that the use of the Rayleigh-Jeans approximation here does not imply that the material is optically thick. It just implies that it follows the same form, but the emission is still modified by the optical depth.

Keep an eye on that last temperature term T. Here, for a blackbody, it is the kinetic temperature. What is it in the more general case? Keep reading to find out!

S constant, LTE, Optically Thin

If $\tau_\nu \ll 1$, we get emission from the background radiation, as well as from along the line of sight. We can make the Taylor expansion substitution $e^{-\tau_\nu} \simeq 1 - \tau_\nu$, so therefore:

$$I_\nu(\tau_\nu) = I_{\nu,0}(\tau_\nu)(1 - \tau_\nu) + B_\nu \tau_\nu \simeq I_{\nu,0}(\tau_\nu) + B_\nu \tau_\nu \tag{1.60}$$

The first term again is the background radiation attenuated by the ISM. The second term is the Planck function modified by the optical depth of the ISM. Notice that we can still have a blackbody-like spectrum even if it is optically thin, although it is modified by the optical depth (which is less than 1). In the case that $\tau_\nu = 0$, we of course only see the background radiation.

In the radio regime, we get

$$T_B = T_{B,0}(1 - \tau_\nu) + T\tau_\nu \simeq T_{B,0} + T\tau_\nu \tag{1.61}$$

S constant, LTE, Optically Thick

If $\tau \gg 1$, $e^{-\tau_\nu} \to 0$, so

$$I_\nu = S_\nu \tag{1.62}$$

This makes sense. If there is a blackbody in our line of sight, we don't see any emission from behind it.

In radio astronomy, $T_B = T$ for optically thick emission, the kinetic temperature of the material (if in LTE).

1.4 Basics of Statistical Mechanics

1.4.1 Maxwell-Boltzmann distribution (partially from Wikipedia)

The Maxwell speed distribution describes particle speeds in idealized gases where the particles move freely inside a stationary container without interacting with one another, except for very brief collisions in which they exchange energy and momentum with each other or with their thermal environment. Applies only to systems in thermodynamic equilibrium. Therefore, we only expect the Maxwell distribution in *dense* gas where collisions are frequent. If collisions are infrequent, energy is not eficiently exchanged and we do not get a Maxwell distribution.

In the Maxwell distribution, the probability of finding a particle with velocity v is:

$$f(v) = \sqrt{\left(\frac{m}{2\pi kT}\right)^3} 4\pi v^2 e^{-\frac{mv^2}{2kT}}, \tag{1.63}$$

where m is the particle mass and kT is the product of Boltzmann's constant and thermodynamic temperature.

We can integrate the MB distribution function to find the mean particle speed,

$$\int_0^\infty v\, f(v)\, dv = \sqrt{\frac{8kT}{\pi m}}, \tag{1.64}$$

or differentiate to find the most likely speed (peak of the function):

$$v_p = \sqrt{\frac{2kT}{m}}\,. \tag{1.65}$$

The root mean square speed is also useful:

$$\sqrt{\langle v^2 \rangle} = \left(\int_0^\infty v^2\, f(v)\, dv \right)^{1/2} = \sqrt{\frac{3kT}{m}} = \sqrt{\frac{3}{2}} v_p\,. \tag{1.66}$$

There is no "right" answer here for the best velocity to choose for all situations.

(from Wikipedia) The Maxwell-Boltzmann distribution applies to the classical ideal gas, which is an idealization of real gases. In real gases, there are various effects (e.g., van der Waals interactions, relativistic speed limits, and quantum exchange interactions) that make their speed distribution sometimes very different from the Maxwell-Boltzmann form. That said, rarefied gases at ordinary temperatures behave very nearly like an ideal gas and the Maxwell speed distribution is an excellent approximation for such gases.

This is in 3D! We actually measure only 1D. Keep this in mind for later.

So, in LTE, all particle species follow a Maxwell-Boltzmann speed distribution that can be characterized by a single (kinetic) temperature T_k. We will assume LTE frequently.

This probability density function gives the probability, per unit speed, of finding the particle with a speed near v. If particles follow a MB distribution, we can characterize them with a single temperature. When does this happen? When frequent collisions are able to thermalize the distribution. Particles at velocity v, move one "mean free path" λ in time t:

$$v = \frac{\lambda}{t} \tag{1.67}$$

The mean free path is

$$\lambda = \frac{1}{n\sigma}\,, \tag{1.68}$$

where n is the particle density and σ is the effective cross section (not necessarily the geometric cross section). Therefore,

$$t \simeq \frac{1}{n\sigma v}\,, \tag{1.69}$$

the particle timescale. This is a useful, although very approximate quantity! This sets the timescale over which a population of particles can thermalize.

The mean free path is related to the optical depth:

$$\tau_\nu = \int \kappa ds = \int n(s)\sigma ds \simeq \frac{s}{\lambda} \tag{1.70}$$

This is telling us something fundamental: when $\tau_\nu = 1$, the photons have traveled one mean free path. Because more photons will have traveled less than one mean free path than more, the mean distance is $< \lambda$.

The temperature of a gas is simply a measure of the kinetic energy of the gas particles. As the temperature goes up, the particles move faster. How fast? Each dimension (degree of freedom) adds $\sim 1/2\,kT$, so the particle energy is roughly $3/2\,kT$. This is approximate, but is roughly the rms speed of a Maxwell-Boltzmann distribution.

1.4.2 Boltzmann Equation

The single most important equation in stat. mech. for us is the Boltzmann Equation:

$$\frac{n_i}{n_j} = \frac{g_i}{g_j} e^{-E_{\rm ij}/k\,T_{\rm ex}}\,, \tag{1.71}$$

where n_i is the density in state i, g_i is the degeneracy of state i, E_{ij} is the energy difference between the two states, and $T_{\rm ex}$ is the "excitation temperature." More on $T_{\rm ex}$ later.

While this equation gives the relative densities between states, we are frequently interested in the fractional density of a given state compared to all states. In such cases, we need to use the "partition function,"

$$Z(T_{\rm ex}) = \sum_i^{\infty} g_i e^{-E_i/k\,T_{\rm ex}} \tag{1.72}$$

So that

$$\frac{n_i}{n_T} = \frac{g_i e^{-E_{\rm ij}/k\,T_{\rm ex}}}{Z(T_{\rm ex})}\,, \tag{1.73}$$

where n_T is the total population in all levels.

The excitation temperature is not a physical temperature! It is instead the temperature at which the Boltzmann equation is satisfied. When is $T_{\rm ex} = T_k$ the kinetic temperature? When collisions are frequent! Assume we have two competing processes: collisions and radiation, and that the kinetic (collision) temperature T_k and radiation temperature (T_R) are different. If the timescale for collisions is closer than the timescale for photon-particle interactions, $T_{\rm ex} \simeq T_k$.

For another example, assume that the population levels are inverted such that the upper level is overpopulated relative to the lower level, then $T_{\rm ex}$ is negative. This is allowed because $T_{\rm ex}$ is not a real temperature. These population inversions can result in masing emission.

It is also worth noting that the excitation temperature only corresponds to the transition between the upper and lower levels. Therefore, each transition can have a different excitation temperature!

Radiation temperature, T_R, is the the equivalent temperature blackbody that would emit the same intensity at the frequency of interest. This is also sometimes called the background temperature, $T_{\rm BG}$ or T_0. In the limit of low frequencies where the RJ limit applies, the brightness and radiation temperatures are the same.

If $T_k = T_R = T_{\rm ex} = T$, the system is in *thermodynamic equilibrium* (TE). This happens when the particle energy distribution follow the Boltzmann equation (and ionization states

follow Saha, see below), the particle velocity distributions follow MB, and the radiation field is a Planck function at temperature T_R.

If $T_k = T_{\rm ex} \neq T_R$, the system is in *local thermodynamic equilibrium*, LTE. LTE is much easier to attain and is commonly assumed. This is often good enough when collisions dominate over radiative processes. "Local" here refers to $\sim$ one mean free path.

In LTE, the changes in temperature must vary slowly, so that at each point in the object of interest we can assume TE. That temperature is that of the particles, which follow a Maxwellian distribution with a single temperature, for all particle species. In other words, the temperature gradient scale must be small compared to the mean free path of the particles.

1.4.3 Saha Equation

The Saha ionization equation relates the ionization state of an element to the temperature and pressure.

For a gas composed of a single atomic species in LTE, only concerning two states (excited and not as excited) the Saha equation is written:

$$\frac{n_{i+1} n_e}{n_i} \simeq 2 \left(\frac{2\pi m_e kT}{h^2} \right)^{3/2} \frac{g_{i+1}}{g_i} \exp\left[-\frac{\Phi_r}{kT} \right] , \tag{1.74}$$

where n_i is the density of atoms in the i-th state of ionization, that is with i electrons removed.
g_i is the degeneracy of states for the i-ions
Φ_r is the energy required to remove i electrons from a neutral atom, creating an i-level ion (the "ionization potential").
n_e is the electron density
m_e is the mass of an electron
T is the temperature of the gas
k_B is the Boltzmann constant
h is Planck's constant .

Hydrogen is particularly simple. The degeneracy for the ground state of hydrogen is 4 (proton spin up, electron up; p up e down; p down e up; p down e down). The degeneracy for the ionized state is 2. We therefore have:

$$\frac{n_{H+} n_e}{n_H} \simeq \left(\frac{2\pi m_e kT}{h^2} \right)^{3/2} \exp\left[-\frac{13.6\,{\rm eV}}{k_B T} \right] , \tag{1.75}$$

Why do we care? Plasmas are a very important component in the ISM. The Saha equation tells us the characteristic temperature for each ionization state of a particular element. But, when is it applicable???

Your book gives ionization potentials for various species and ionization states in Appendix D. **For the common elements in the ISM, these ionization potentials are incredibly useful to know!** As atoms lose electrons, the remaining electrons require more and more energy to become ionized.

1.4.4 Review: Temperatures

Already, we have seen many different definitions of temperatures. Let's review these all in one place. Some of these will be new.

- Excitation Temperature ($T_{\rm ex}$): The value required for the Boltzmann equation to return the observed level populations. This is different for each transition (pair of levels)!
- Spin Temperature (T_s): The excitation temperature for the 21 cm H I line.
- Radiation Temperature (T_R): The temperature that characterizes a blackbody representative of the radiation field. For example, $T_R = 2.7$ K for the CMB.
- Kinetic Temperature (T_K): The temperature that characterizes the distribution of particle energies. Typically the rms speed of a Maxwell-Botzmann distribution is the most useful, to $T_K = (3kT/m)^{0.5}$.
- Brightness Temperature (T_B): The temperature a black body in thermal equilibrium with its surroundings would have to be to duplicate the observed intensity of a grey body object at a frequency.

And we have two simplifying regimes:

- Thermodynamic Equilibrium (TE): The particle energy distribution follows the Boltzmann equation (and ionization states follow Saha), the particle velocity distributions follow MB, and the radiation field is a Planck function at temperature T_R. In this case, IF $T_k = T_R = T_{\rm ex}$.
- Local Thermodynamic Equilibrium (LTE): TE applies over one particle mean free path, or collisions dominate over radiative processes. In this case, $T_k = T_{\rm ex} \neq T_R$.

1.4.5 Lasers and Masers!

"Laser" stands for Light Amplification by Stimulated Emission of Radiation. Masers are lasers in the microwave (radio) regime. Most such detected emission is from masers rather than lasers, so we will use this term throughout. Masers are caused by stimulated emission, covered earlier. We say that an object that emits maser emission is "masing." Emission need not be completely masing or non-masing - it is usually a mix in fact. We will cover masers in more depth later, but masers have some interesting terms relating to radiative transfer.

In a maser, we have a "population inversion," where the upper energy levels of an atom or

molecule are over populated compared to the lower levels for the object temperature. In such cases, impinging radiation stimulates the electronic transition from the over-populated upper state to the lower state. In such cases, the attenuation is negative, the excitation temperature is negative, and the optical depth is negative.

1.5 Spectroscopy

The first broad ISM topic for this course is gas (excluding ionized gas), its physics, how we detect it, and what we know from it. Gas is the most massive barionic component of the Galactic ISM. It also has the benefit of interesting, relatively simple physics (at least compared to dust). We will work our way from the ubiquitous atomic H I through the dense gas that is forming stars. Before we can start that though, we need a bit of background on radiation. This review will cover most everything that you will need, but I expect that you have some basic knowledge already.

Observations of gas in the ISM primarily consist of spectra, the intensity as a function of frequency (or wavelength or velocity, etc). We can divide a spectrum into **"line"** and **"continuum"** components. A spectral line is a decrement or excess intensity on top of the continuum [draw figure of spectral line]. Spectral lines are produced from discrete (quantized) transitions, most simply by electrons through electric dipole radiation, although many more emission mechanisms are possible. Continuous radiation can come from a variety of different emission mechanisms, and we will cover those later in the course when we talk about ionized gas.

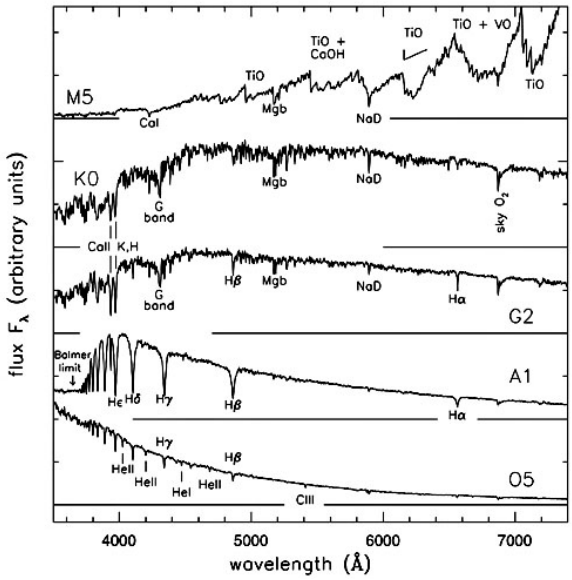


Figure 1.13: Spectra for stars of various spectral types, showing absorption lines on top of the continuum.

There are two broad types of line spectra:
Absorption lines (decreases in intensity relative to nearby continuum) are produced when

a cold gas is between a continuous spectrum source and the detector. Atoms in the cold gas absorb photons from the continuous spectrum source. The photons are then emitted in random directions, removing intensity along the line of sight. What is a continuous spectrum source? Stars and quasars are pretty close (especially quasars). Most stellar spectra are absorption line spectra, however, because the outer layers of the photosphere are less dense and colder than the core.

Emission lines at discrete wavelengths are produced when the detector sees photons emitted directly from a (hot) gas. The wavelengths of the emission lines are due to electronic transitions within the atoms and are therefore unique to each atom.

1.5.1 Kirchoff's Laws

The spectrum you observe depends on the density (the optical depth) of the object, and the viewing direction. Observing the same object from a different direction will give you a different signal. **Kirchoff's Laws** tell us how to interpret the spectra we observe. There are multiple sets of Kirchoff's Laws, so it is safe to assume that Kirchoff was wicked smart and interesting at parties.

Kirchoff's three laws of spectra are:

- A dense object produces light with a continuous (blackbody) spectrum. Kirchhoff also coined the term blackbody radiation because he was a show-off. You emit blackbody radiation, with a peak in the infrared.
- A hot diffuse gas produces an emission line spectrum due to electronic transitions within the gas. Fluorescent lights are a good example.
- A hot dense object surrounded by a cool tenuous gas (i.e., cooler than the hot object) produces an absorption line spectrum. The absorption lines are at exactly the same wavelengths as the emission lines for a given element, and are also due to electronic transitions.

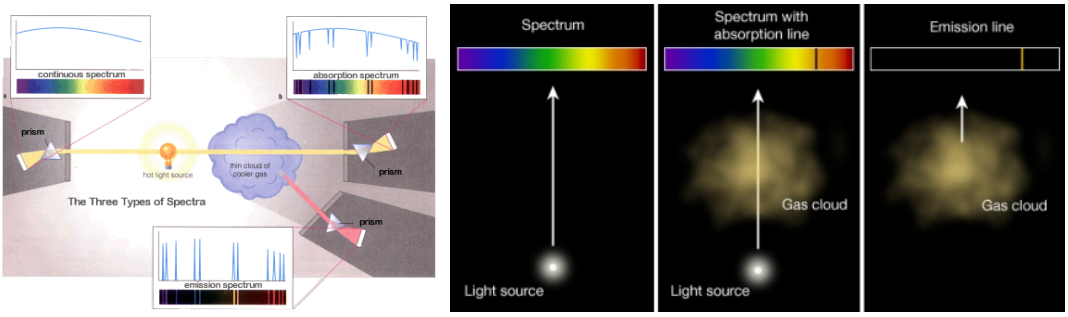


Figure 1.14: Kirchoff's Laws

1.5.2 Spontaneous Emission, Stimulated Emission, and Absorption (Draine Chapter 6)

The quantization of energy levels allows us to create some general principles of emission and absorption. There are three separate classes of transitions: Spontaneous Emission, Stimulated Emission, and Absorption. These transitions need not be electric dipole, but rather can describe any transitions between quantized energy levels, for atoms, ions, molecules, or dust grains.

These transitions each have a corresponding **"Einstein Coefficient."** Einstein was also wicked smart. The Einstein coefficients help us to relate the transition probabilities to the particle and energy densities of the medium.

Absorption

An absorber X in level ℓ can absorb a photon with energy $h\nu$:

$$X_\ell + h\nu \rightarrow X_\mathrm{u} \tag{1.76}$$

This only happens of course if the photon energy $h\nu$ corresponds to the difference in energy levels of the object in question.

If we have the volume density n_ℓ in the lower state, the volume absorption rate is:

$$\left(\frac{dn_\mathrm{u}}{dt}\right)_{\ell\rightarrow\mathrm{u}} = -\left(\frac{dn_\ell}{dt}\right)_{\ell\rightarrow\mathrm{u}} = B_{\ell\mathrm{u}} n_\ell u_\nu \,, \tag{1.77}$$

where u_ν is the spectral energy density at the correct frequency. The constant $B_{\ell\mathrm{u}}$ is the Einstein B coefficient for the transition $\ell \rightarrow$ u. We can see that the units of the Einstein B coefficient must be probability per unit time per unit spectral energy density [$\mathrm{erg\,s^{-1}\,cm^3\,Hz^{-1}}$] of the radiation field.

We can write the above equations because transitions out of level ℓ must populate level u, and vice-versa. We will use this fact repeatedly.

Spontaneous Emission

Spontaneous emission is the process by which an electron "spontaneously" (i.e., without any outside influence) decays from a higher energy level to a lower one:

$$X_\mathrm{u} \rightarrow X_\ell + h\nu \,. \tag{1.78}$$

Similar to the above, we can say

$$-\left(\frac{dn_\mathrm{u}}{dt}\right)_{\mathrm{u}\rightarrow\ell} = \left(\frac{dn_\ell}{dt}\right)_{\mathrm{u}\rightarrow\ell} = A_{\mathrm{u}\ell} n_\mathrm{u} \,, \tag{1.79}$$

where $A_{\mathrm{u}\ell}$ is the Einstein A coefficient for spontaneous emission.

The units for the Einstein A coefficient are simply per unit time. Spontaneous emission is independent of the radiation field. Therefore, the lifetime in the excited state is just A^{-1}. If the Einstein A is high, the lifetime must be low and the transitions happen rapidly. If the Einstein A is low, the lifetime is high, and the transitions happen at a slower rate. Therefore, all else being equal, the strongest transitions have high Einstein A's. We will develop this concept further later.

The Einstein A coefficient has a particularly simple form for electric dipole radiation:

$$A_{\mathrm{u}\ell} = \frac{64\pi^4}{3hc^3}\nu_{\mathrm{u}\ell}^3|\mu_{\mathrm{u}\ell}|^2\,, \tag{1.80}$$

where $\mu_{\mathrm{u}\ell}$ is the mean electric dipole moment. What Equation 2.5 shows is that the Einstein A increases rapidly as frequency increases. [What does this tell you about the line intensity?] Transitions are much faster, and the line intensity therefore stronger, in the optical compared to the radio!

Stimulated Emission (from Wikipedia)

Stimulated emission (also known as induced emission) is the process by which an electron is induced to jump from a higher energy level to a lower one by the presence of electromagnetic radiation at (or near) the frequency of the transition. This is the basis of lasers and masers. From the thermodynamic viewpoint, this process must be regarded as negative absorption. We can describe the process:

$$X_{\mathrm{u}} + h\nu \rightarrow X_\ell + 2h\nu\,. \tag{1.81}$$

$$-\left(\frac{dn_{\mathrm{u}}}{dt}\right)_{\mathrm{u}\rightarrow\ell} = \left(\frac{dn_\ell}{dt}\right)_{\mathrm{u}\rightarrow\ell} = B_{\mathrm{u}\ell}n_{\mathrm{u}}u_\nu\,, \tag{1.82}$$

where $B_{\mathrm{u}\ell}$ is the Einstein coefficient for stimulated emission (note that negative absorption is entirely appropriate given the nomenclature). Both Einstein B coefficients have the same units.

In stimulated emission a photon induces the electronic transition. The resultant photons have exactly the same energy, phase, polarization and direction of propagation. Such radiation is called "coherent."

Putting these relations together, we have

$$\frac{dn_u}{dt} = \left(\frac{dn_u}{dt}\right)_{\ell\rightarrow\mathrm{u}} + \left(\frac{dn_u}{dt}\right)_{\mathrm{u}\rightarrow\ell} = n_\ell B_{\ell\mathrm{u}}u_\nu - n_u\left(A_{\mathrm{u}\ell} + B_{\mathrm{u}\ell}u_\nu\right)\,. \tag{1.83}$$

The Einstein coefficients are not independent of each other. Absorption of a photon can lead to emission at some later time. Similarly, emission of a photon, either by spontaneous or stimulated emission, can lead to absorption.

In TE, the net exchange between any two levels will be balanced and $\frac{dn_u}{dt} = 0$. This is known as "detailed balance." From the above expression, we can therefore write:

$$0 = A_{\mathrm{u}\ell}n_{\mathrm{u}} + B_{\mathrm{u}\ell}n_{\mathrm{u}}u_\nu - B_{\ell\mathrm{u}}n_\ell u_\nu\,. \tag{1.84}$$

Draine shows how these coefficients must be related in the case that u_ν is the energy density of a blackbody[1] and n_u and n_ℓ are related with the Boltzmann function (in TE, as above):

$$B_{\mathrm{u}\ell} = \frac{c^3}{8\pi h\nu^3} A_{\mathrm{u}\ell} \tag{1.85}$$

or

$$B_{\ell\mathrm{u}} = \frac{g_\mathrm{u}}{g_\ell} B_{\mathrm{u}\ell} = \frac{g_\mathrm{u}}{g_\ell} \frac{c^3}{8\pi h\nu^3} A_{\mathrm{u}\ell}\,. \tag{1.86}$$

The strength of the stimulated emission and absorption are both determined by stimulated emission and the ratio of the state degeneracies.

1.5.3 A two-level system

We have a definition for the excitation temperature, but how can we actually calculate it? We need to find the population ratio between two states. The simplest possible exercise here is a two-level state, again with levels u and ℓ.

Two types of processes can induce transitions between the two states: radiative and collisional. We can induce upward transitions with: Collisional excitation or absorption. We can induce downward transitions with collisional de-excitation, spontaneous emission, or stimulated emission.

The excitation temperature is determined by the level population ratio. We can use detailed balance, derived above, but including collisions. In detailed balance, although not in strict TE, the number of transitions out of a level must equal the number of transitions into that level.

The transition rates are:
(1) Collisional excitation rate: $C_{\ell\mathrm{u}} n_\ell$
(2) Absorption rate: $B_{\ell\mathrm{u}} n_\ell u_\nu$
(3) Collisional de-excitation rate: $C_{\mathrm{u}\ell} n_\mathrm{u}$
(4) Spontaneous emission rate: $A_{\mathrm{u}\ell} n_\mathrm{u}$
(5) Stimulated emission rate: $B_{\mathrm{u}\ell} n_\mathrm{u} u_\nu$

C here is the collision rate coefficient. The collision rate [s^{-1}] is just the particle velocity divided by the mfp,

$$C = n\sigma v\,. \tag{1.87}$$

In detailed balance, all excitations must be balanced by de-excitations, so

$$n_\ell (C_{\ell\mathrm{u}} + B_{\ell\mathrm{u}} u_\nu) = n_\mathrm{u} (C_{\mathrm{u}\ell} + B_{\mathrm{u}\ell} u_\nu + A_{\mathrm{u}\ell}) \tag{1.88}$$

[1] $u_\nu(T) = \frac{4\pi}{c} B_\nu(T)$. The energy density has units of energy per volume per frequency unit.

I will leave it as an exercise to show [hint: think of limit of Equation 1.88 when collisions dominate, assume collisional and radiative processes are completely independent, and use

$$u_\nu = \frac{4\pi}{c}\frac{2h\nu^3}{c^2}\frac{1}{e^{h\nu/kT_R}-1}\,, \tag{1.89}$$

to find

$$e^{-h\nu/k\,T_{\rm ex}} = \frac{e^{-h\nu/kT_k}C_{{\rm u}\ell} + \frac{1}{e^{h\nu/kT_R}-1}A_{{\rm u}\ell}}{C_{{\rm u}\ell} + \left(\frac{1}{e^{h\nu/kT_R}-1}+1\right)A_{{\rm u}\ell}} \tag{1.90}$$

Equation 1.90 has two limits:
(1) Collisions dominate, so $C_{{\rm u}\ell} \gg A_{{\rm u}\ell}$. Then

$$e^{-h\nu/k\,T_{\rm ex}} \simeq \frac{e^{-h\nu/kT_{\rm k}}C_{{\rm u}\ell}}{C_{{\rm u}\ell}} = e^{-h\nu/kT_{\rm k}} \tag{1.91}$$

So $T_{\rm ex} = T_k$. When collisions dominate, the level populations are in equilibrium with the gas kinematic processes. This is the case in high densities.

(2) Radiation dominates, so $A_{{\rm u}\ell} \gg C_{{\rm u}\ell}$. Then

$$e^{-h\nu/k\,T_{\rm ex}} \simeq \frac{\frac{1}{e^{h\nu/kT_{\rm R}}-1}A_{{\rm u}\ell}}{\left(\frac{1}{e^{h\nu/kT_R}-1}+1\right)A_{{\rm u}\ell}} = e^{-h\nu/kT_{\rm R}} \tag{1.92}$$

So $T_{\rm ex} = T_R$. When radiation dominates (at low densities), the level populations are in equilibrium with the radiation field.

Both of these can be in LTE! Remember: LTE just requires that the changes in temperature are slowly varying.

1.5.4 Critical Density

The transition from $T_{\rm ex} = T_R$ to $T_{\rm ex} = T_K$ happens when collisional excitations balance spontaneous emission:

$$C = n\sigma v = A_{{\rm u}\ell}\,. \tag{1.93}$$

Very approximately, this happens at a critical density $n_{\rm crit}$ of

$$n_{\rm crit} \simeq \frac{A_{{\rm u}\ell}}{\sigma v}\,. \tag{1.94}$$

Equation 1.94 assumes optically thin emission, which is frequently not appropriate. In reality, often there is **radiative trapping**, which makes a particular transition optically thick to its own radiation. Imagine background molecules radiating and that background radiation exciting foreground molecules (instead of collisions). We will deal with this more complicated scenario for CO in our chapter on molecules.

We can think of the critical density as the density required to excite the atom or molecule into a specific energy state. Critical densities are especially important for molecular line observations. The (optically thin) critical density for CO $J = 1 - 0$ transitions is about 700 cm^{-3}, while it is about 10^3 cm^{-3} for NH_3.

Because the critical density is so important for interpreting spectral line emission, we have some specialized vacabulary for discussing it. If $T_{\rm ex} < T_K$, we say the level populations are "subthermally excited." When $T_{\rm ex} \simeq T_K$, we say that the level populations are "thermalized."

Interestingly, in the optical T_R can be much higher than $T_{\rm ex}$. Therefore, at densities less than the critical density, optical emission lines can be bright if T_R is larger than $T_{\rm ex}$, but radio emission lines will be faint.

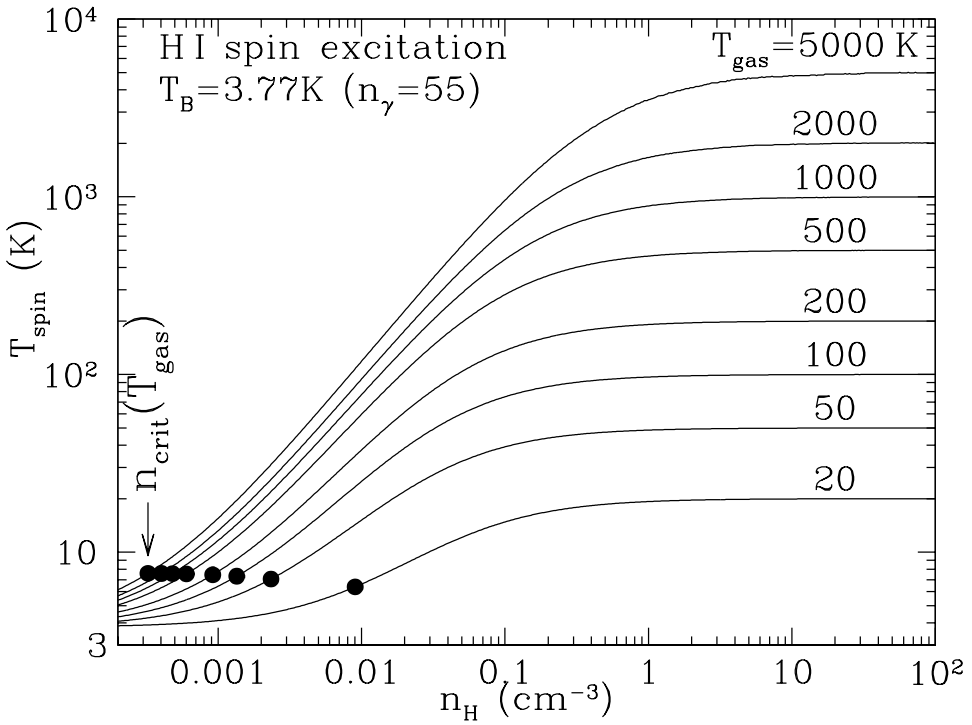


Figure 1.15: H I excitation temperature ("spin temperature") as a function of density. We see that at low densities, collisions are unimportant and $T_s \simeq T_R$ (T_R written as T_B here). At high densities, collisions are important and $T_s \simeq T_k$. The critical densities occur at spin temperatures between T_R and T_k.

1.5.5 Revisiting $T_{\rm ex}$

The excitation temperature is in between the radiation temperature and the kinetic temperature. We can solve Equation 1.90 to find

$$\frac{1}{T_{\rm ex}} = \frac{\frac{1}{T_k} + \frac{A_{u\ell}}{C_{u\ell}}\frac{T_R}{T_0}\frac{1}{T_R}}{1 + \frac{A_{u\ell}}{C_{u\ell}}\frac{T_R}{T_0}} \tag{1.95}$$

where $T_0 = h\nu/k$ [Rohlfs & Wilson, 11.4.].

Note that $T_{\rm ex}$ is *not* the average between T_k and T_R. Written this way, the excitation temperature is a harmonic mean between the radiation and kinetic temperatures weighted by the relative collisional and radiative de-excitation terms. The weighting factor in the harmonic mean is the ratio of the density to the critical density. However, note that this factor is multiplied by an additional factor of T_R/T_0. If stimulated emission is important, $T_0 \ll T_R$, and the density must be much larger than $n_{\rm crit}$ to produce visible emission lines.

It is worth rewriting this equation once again, since it can be solved to give

$$T_{\rm ex} = T_K \frac{(T_R + T_0)A_{u\ell} + C_{u\ell}T_0}{A_{u\ell}T_K + C_{u\ell}T_0} \tag{1.96}$$

At low radio frequencies, $(T_R+T_0) \simeq T_R$. We see again that if $A_{u\ell} \gg C_{u\ell},\ T_{\rm ex} = T_R+T_0 \simeq T_R$, and $C_{u\ell} \gg A_{u\ell},\ T_{\rm ex} = T_K$.

The location of $n_{\rm crit}$ along the varying portion of the curve depends on the importance of stimulated emission. Or, the difference between $T_{\rm ex}$ at $n = n_{\rm crit}$ and T_k tells you about stimulated emision. If there is a large difference, stimulated emission is important (low frequencies).

[Pogge notes] For example, take the excitation temperature as a function of density for two molecular lines: the CS J=3-2 147 GHz line and the NH_3 (J,K)=(1,1) 23 GHz line. This CS line has a critical density of 1.5×10^6 cm^{-3}, and reaches an excitation temperature of $0.9T_k$ at $n > 10^6$ cm^{-3}. However, the NH_3 line has a critical density of 2×10^3 cm^{-3}, but the excitation temperature does not approach the kinetic temperature until $n > 10^5$ cm^{-3}, nearly 3 orders of magnitude larger. The reason is the greater importance of stimulated emission at low frequencies compared to at high frequencies. In general, if $T_0 = h\nu/k \ll T_k$, the density must be much larger than the critical density in order for the line to be visible. As such, the density must be very large compared to the critical density to thermalize lines at centimeter wavelengths (like the NH_3 transition noted above), while at millimeter and sub-millimeter wavelengths, the lines of species like CO and CS are essentially thermalized at or near their critical densities. By the time you get to Infrared and Visible wavelengths stimulated emission becomes negligible and all collisionally excited lines thermalize at $n_{\rm crit}$.

We will revisit these concepts in our chapters on atomic and molecular gas.

1.5.6 Emission and Absorption Coefficients (Draine 7.3)

The meaning of j_ν and κ_ν are a bit difficult to discern, so let's unpack them a bit. I will revisit this discussion frequently, so don't be too worried if it doesn't make sense at first.

We can write a general form for the emissivity in the case of isotropic radiation as

$$j_\nu = n_{\rm u} \frac{A_{{\rm u}\ell}}{4\pi} h\nu_{{\rm u}\ell} \phi_\nu \tag{1.97}$$

where ϕ_ν is the normalized line profile, $\int \phi_\nu \, d\nu = 1$ (more on this later). This line profile tells us about the shape of the spectral line (usually assumed to be Gaussian), as a function of frequency. What this is saying is that the emissivity (units $\mathrm{erg\,cm^{-1}\,sr^{-1}\,s^{-1}}$) is proportional to the Einstein A, the frequency, and the density in the upper state. At line center, ϕ_ν is at maximum, and so is j_ν; both fall off after that. We see that as $A_{{\rm u}\ell}$ increases, j_ν increases – as transitions become more frequent we get more total power.

Note that the presence of stimulated emission alone does not mean that it is dominant, and the source is a maser! Stimulated emission is always present at some level. Only when it is dominant do we consider the source to be a maser.

The absorption coefficient is just absorption minus stimuated emission ("negative absorption"). Pure absorption looks like:

$$\kappa_{\nu,absorption} = \frac{h\nu}{4\pi} B_{{\rm u}\ell} n_\ell \,, \tag{1.98}$$

whereas stimuated emission looks like

$$\kappa_{\nu,st.emission} \propto B_{\ell{\rm u}} n_{\rm u} \tag{1.99}$$

and using the Boltzmann equation and our earlier relationship between Einstein B coefficients we find

$$\kappa_{\nu,st.emission} \propto B_{{\rm u}\ell} e^{-h\nu_{{\rm u}\ell}/k\,T_{\rm ex}} \,. \tag{1.100}$$

Using our relationship between $B_{{\rm u}\ell}$ and $A_{{\rm u}\ell}$ in Equation 1.85, the total linear absorption coefficient is just the absorption minus the stimulated emission ("negative absorption"). Draine does a nice derivation to get:

$$\kappa_\nu = \kappa_{\nu,absorption} - \kappa_{\nu,st.emission} = n_\ell \frac{g_{\rm u}}{g_\ell} \frac{A_{{\rm u}\ell}}{8\pi} \lambda^2_{{\rm u}\ell} \phi_\nu \left[1 - e^{-h\nu_{{\rm u}\ell}/k\,T_{\rm ex}}\right] \,. \tag{1.101}$$

The terms before the bracket should all be familiar. The "1" inside the brackets refers to absorption, which depends on the density in the lower state, the Einstein A, the ratio of the degeneracies, and the frequency of the transition. The second term inside the brackets is due to stimulated emission. If $e^{-h\nu_{{\rm u}\ell}/k\,T_{\rm ex}} \simeq 1$, stimulated emission is important! When does this happen? When $-h\nu_{{\rm u}\ell}/k\,T_{\rm ex} \lesssim 0$. The stimulated emission term is larger when the frequency of the transition is low, or when $T_{\rm ex}$ is large, or when $T_{\rm ex}$ is negative. When $T_{\rm ex}$ is negative, κ_ν is also negative. This is a maser. [Draine likes some terms of ν and some of λ for some reason.]

Again applying Boltzmann, we can find that the source function is:

$$S_\nu = j_\nu/\kappa_\nu = \frac{\frac{n_{\rm u} h\nu}{4\pi} A_{{\rm u}\ell}}{\frac{c^2}{8\pi\nu^2} A_{{\rm u}\ell} n_{\rm u} \left[e^{h\nu/k\,T_{\rm ex}} - 1\right]} \tag{1.102}$$

which reduces to

$$S_\nu = \frac{2h\nu^3}{c^2} \frac{1}{e^{h\nu/k\,T_{\rm ex}} - 1}\,. \tag{1.103}$$

The source function of a spectral line is the Planck function at a temperature of $T_{\rm ex}$. This is not the intensity, but just the source function. The intensity will be the source function modified by the optical depth.

Let's remember our solution to the equation of radiative transfer in the case that the source function is constant

$$I_\nu(\tau_\nu) = I_{\nu,0} e^{-\tau_\nu} + S_\nu(1 - e^{-\tau_\nu})\,, \tag{1.104}$$

or in the radio regime

$$T = T_0 e^{-\tau_\nu} + T_{\rm ex}(1 - e^{-\tau_\nu})\,. \tag{1.105}$$

Note that our T that replaced the source function is actually the excitation temperature via Equation 1.103. In LTE, $T_{\rm ex} = T_k$ and

$$T = T_0 e^{-\tau_\nu} + T_k(1 - e^{-\tau_\nu})\,. \tag{1.106}$$

1.5.7 Line Broadening

Emission and absorption lines are not delta functions, they are broadened by some mechamisms. What can broaden spectral lines?
1) Thermal doppler motion due to gas particles at a given temperature, "Doppler broadening." Not all particle speeds will be the same of course. As the temperature increases, the range of speeds does too.
2) "Turbulent broadening" due to the fact that there are bulk motions within any gas.
3) Natural broadening due to the fact that the energy "level" is not a single value, and
4) Pressure or collisional broadening, which changes the energy levels.

Each of these mechanisms has an associated line shape. Doppler broadening is Gaussian, and turbulent is usually assumed to be Gaussian as well. Natural and Collisional broadening are "Lorenzian," which is like a Gaussian but with much larger "wings." All four processes operate at the same time, resulting in a "Voigt" profile with a Gaussian core and Lorentzian wings. Since the Lorentzian wings are at low intensity, usually something best approximated by a Gaussian is observed.

Gaussians are magical functions. A normalized Gaussian takes the form of

$$\phi(\nu) = \frac{1}{\sigma\sqrt{2\pi}} {\rm e}^{-\left(\frac{(\nu-\nu_0)^2}{2\sigma^2}\right)}\,, \tag{1.107}$$

where ν_0 is usually the line center and σ is the one-dimensional velocity dispersion. You can see that the line will be of maximum intensity when $\nu = \nu_0$ at line center, then $\phi(\nu_0) = \frac{1}{\sigma\sqrt{\pi}}$.

We can also define the full-width at half-maximum (FWHM) as

$$FWHM = \sqrt{8\ln 2}\,\sigma = 2.355\sigma\,. \tag{1.108}$$

For thermal gas with particles of mass m following a MB distribution,

$$\sigma = \left(\frac{kT}{m}\right)^{1/2} = 9.12\left(\frac{T_4}{m/\mathrm{amu}}\right)^{1/2}\ \mathrm{km\,s^{-1}}\,. \tag{1.109}$$

The constant goes to 21.47 $\mathrm{km\,s^{-1}}$ for the FWHM.

The thermal line width increases with increasing temperature and decreases with increasing mass.

Gaussians have the amazing property that the area under the curve is approximately the FWHM times the peak. They are also their own Fourrier transform pair, and a Gaussian convolved with a Gaussian leads to another Gaussian. This last point is very important in astronomy. Frequently, the spectral response of your instrument can be assumed to be Gaussian, and the source or spectral line Gaussian as well. Therefore, you will observe a Gaussian.

Why do we get Gaussians? We had the MB velocity distribution before:

$$f(v) = \left(\frac{m}{2\pi kT}\right)^{3/2} 4\pi v^2 e^{-mv^2/2kT}\,. \tag{1.110}$$

What you have probably never seen is that in 1D the $4\pi v^2$ term goes away. This term arises in 3D due to the density of velocity states available (see Hyperphysics site). This is a Gaussian! Looking at Equation 1.107, the original exponent numerator was $-(\nu - \nu_0)$ and is now $-v^2$. The original exponent denominator was $2\sigma^2$ and is now $2kT/m$. We can define:

$$\sigma_v = \left(\frac{kT}{m}\right)^{1/2} = 9.12\left(\frac{T_4}{m/\mathrm{amu}}\right)^{1/2}\mathrm{km/s}\,, \tag{1.111}$$

where T_4 is the temperature in units of $10^4\,\mathrm{K}$. The FWHM is 2.355σ, or

$$FWHM = 2.355\sigma = 21.47\left(\frac{T_4}{m/\mathrm{amu}}\right)^{1/2}\mathrm{km/s}\,. \tag{1.112}$$

This leads to:

$$f(v) = \frac{1}{\sqrt{2\pi}}\frac{1}{\sigma_v}e^{-u^2/2\sigma_v^2}\,. \tag{1.113}$$

You can see how the velocity dispersion σ_v goes into the Gaussian.

We can assume that turbulence adds another Gaussian term, and we can add the thermal and turbulent components in quadrature, which leads to

$$\sigma = \left(\frac{kT}{m} + v_{\rm turb}^2 \right)^{1/2} , \tag{1.114}$$

We often don't know $v_{\rm turb}$, but given a spectral line we can determine it from the linewidth if we can estimate the temperature.

The natural width arises due to the uncertainty principle: $\Delta E \Delta t \sim \hbar$. Here, $\Delta t = A_{{\rm u}\ell}^{-1}$, so short-lived states have large uncertainties in energy. Because for electric dipole radiation $A_{{\rm u}\ell} \propto \nu^3$, $\Delta E \propto \nu^3$ and natural broadening is important at high frequencies. It gives rise to a Lorentzian profile function

$$\phi(\nu) = \frac{\gamma}{4\pi^2} \frac{1}{(\nu - \nu_0)^2 + (\gamma/4\pi)^2} , \tag{1.115}$$

where γ is a constant for each species related to the spontaneous decay rates

$$\gamma = \Sigma A_{{\rm u}\ell}. \tag{1.116}$$

Like natural broadening, collisional broadening distorts the energy levels, leading to an additional Lorentzian term that can be combined with natural broadening:

$$\phi(\nu) = \frac{\Gamma}{4\pi^2} \frac{1}{(\nu - \nu_0) + (\Gamma/4\pi)^2} , \tag{1.117}$$

where $\Gamma = \gamma + 3\nu_{\rm col}$, and $\nu_{\rm col}$ is the collision frequency, $\nu_{\rm col} = n\sigma v$.

In the radio, Doppler and turbulent broadening are typically the largest. At high frequencies (X-ray), natural braodening can make a large contribution.

1.5.8 Characterizing Spectral Lines

It is often preferable to fit a line profile to a spectral line in order to characterize its emission. An alternative method, called the "equivalent width," is insensitive to the exact profile. The equivalent width is the width, in wavelength or frequency units of a rectangular area equal to that of the spectral line:

$$W_\lambda = \int (1 - F_\lambda / F_0) d\lambda \tag{1.118}$$

The intensity of a spectral line changes with the optical depth at line center. This is usually parameterized as the change in the optical depth, or the change in "column density," the integral of the number density over the pathlength.

Let's assume we have an absorption line and no source function:

$$\frac{I_\nu}{I_{\nu,0}} = e^{-\tau_\nu} \tag{1.119}$$

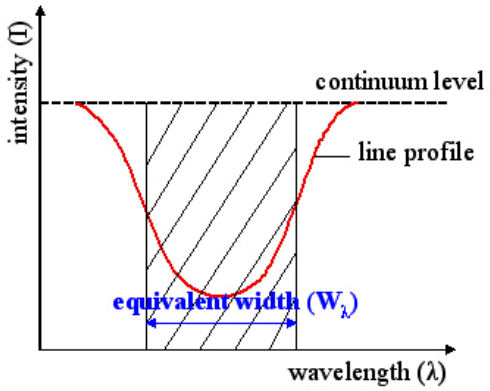


Figure 1.16: Equivalent width.

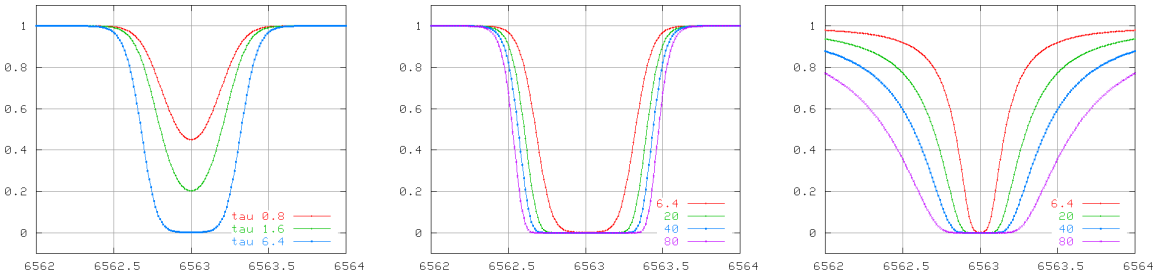


Figure 1.17: Line saturation for various optical depths (http://spiff.rit.edu/classes/phys440/lectures/curve/curve.html).

Initially, the more absorbing atoms there are, the stronger the absorption line. If we keep adding absorbers to the path, however, eventually the line will saturate. A saturated line means that at line center (the most probable speed in a MB distribution) no more intenisity can be added. After it saturates, the equivalent width grows slowly, because there are few atoms with the requisite speeds, until the growth of the line wings becomes important.

The column density is a fundamental quantity in astronomy, since it is directly related to what we measure. Since $\kappa_\nu \propto n$ and $\tau_\nu = \int \kappa_\nu \, ds$, $\tau \propto \int n \, ds = N$. This is the number of particles along a $1\,\mathrm{cm}^2$ cylinder path.

This leads to three important regimes:
1) The "Linear" regime ($\tau \lesssim 5$) where the equivalent width W is proportional to the column density, $W \propto N$.
2) The "saturated" regime where $W \propto \sqrt{\ln N}$, and
3) The "damping" regime where collisional broadening takes over, $W \propto \sqrt{N}$.

[Reading for next time, Draine Chapter 8]

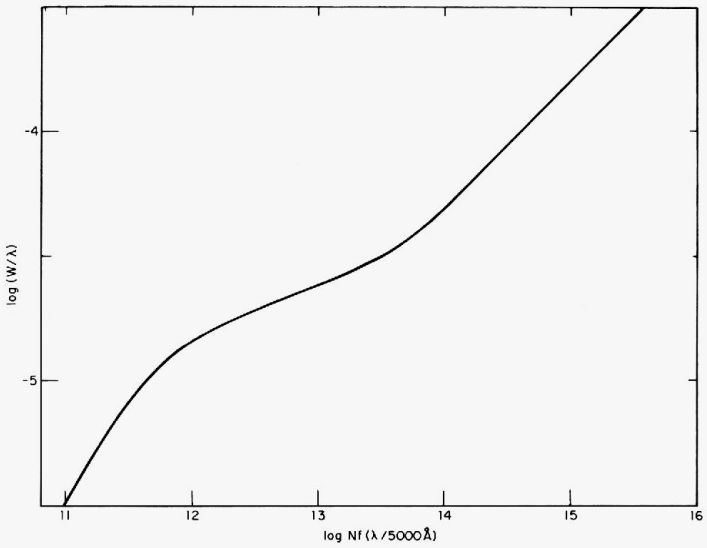


Figure 9.22 A general curve of growth for the Sun. (Figure from Aller, *Atoms, Stars, and Nebulae*, Revised Edition, Harvard University Press, Cambridge, MA, 1971.)

Figure 1.18: The curve of growth. The x-axis is parameterized in terms of the oscillator strength, which is a parameter unique for each transition.

Chapter 2

Basics of HI

The ISM is primarily hydrogen (60% of all gas), and the hydrogen is primarily in its atomic state. We are fortunate to be able to detect an emission line from Hydrogen: the 21 cm line. This is the most important line in all of astrophysics, because it allows us to discern the total mass of galaxies, and because it is so bright that we can detect it across much of the Universe.

The 21 cm H I line is caused by a hyperfine spin-flip transition. *This is not an "allowed" transition because it violates the electric dipole selection rules! It is magnetic dipole rather than electric dipole radiation.* The electron spin can either be aligned (high energy configuration) or anti-aligned (low energy configuration) with that of the proton. The electron spin will flip occasionally, in the transition from the $S = 1$ level to the $S = 0$ level. The difference in energy is only 5.87×10^{-6} eV.

Just like the electrons, the nuclei have an intrinsic spin angular momentum, which we denote $\vec{I}$. Summing over all the nucleons:

$$\vec{I} = \sum_i \vec{s_i}\,, \tag{2.1}$$

where s_i is the spin of the i^{th} nucleon.

If $\vec{I}$ is non-zero, the molecule will have an intrinsic magnetic dipole moment

$$\vec{\mu_N} = g\frac{e}{2Mc}\vec{I}\,, \tag{2.2}$$

where g is the "Landé g factor and M is the nuclear mass. The Landé g factor is

$$g = g(J, L, S) = 1 + \frac{J(J+1) + S(S+1) - L(L+1)}{2J(J+1)} \tag{2.3}$$

The dipole moment of the nuclear produces a magnetic field that interacts with the dipole moment of the electrons. This interactions adds additional internal energy to the atom and splits the levels into "sublevels." These sublevels are known as "hyperfine states."

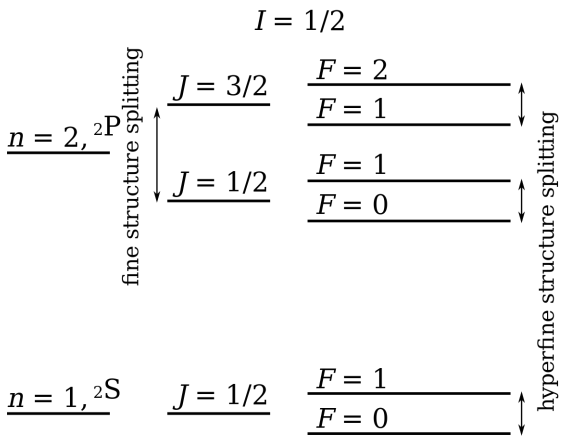


Figure 2.1: Fine-structure and hyperfine splitting for hydrogen (from Wikipedia).

We use a new quantum number F

$$\vec{F} = |\vec{I} + \vec{J}| \tag{2.4}$$

It has values as you would expect. Say you have a $^3\mathrm{D}_3$ atom in the $J = 3$ state with $I = 2$, then $F = 1, 2, 3, 4, 5$. In other words, the $^3\mathrm{D}_3$ energy has 5 hyperfine splittings.

When do we get hyperfine splitting? When there is a non-zero nuclear spin. Obviously, odd numbers of nucleons give non-zero nuclear spins. Even numbers of nucleons usually, but don't always, have zero nuclear spin.

The most famous hyperfine transition is that of H I of course. For H I, $I = 1/2$ and $J = 1/2$, so $F = 0, 1$. The 21 cm line transition is from $F = 1 \to 0$.

The H I line emits via magnetic dipole radiation. By analogy with the Einstein A for electric dipole radiation,

$$A_{\mathrm{u}\ell} = \frac{64\pi^4}{3hc^3}\nu_{\mathrm{u}\ell}^3 |\mu_{\mathrm{u}\ell}|^2 \,, \tag{2.5}$$

the Einstein A is

$$A_{\mathrm{u}\ell} = \frac{64\pi^4}{3hc^3}\nu_{\mathrm{u}\ell}^3 |\mu_{\mathrm{u}\ell}^*|^2 \,, \tag{2.6}$$

where $\mu_{u\ell}^*$ is the mean magnetic dipole moment for H I in the ground electronic state. With a value for this constant, the Einstein A for the H I line is $A_{\mathrm{u}\ell} \simeq 2.85 \times 10^{-15}\,\mathrm{s}^{-1}$. The radiative lifetime is therefore about 11 Myr. We only see H I emission because there is so much H I.

The "spin temperature" is an important parameter in the study of H I gas, as it can be readily derived from observations (with some assumptions). For H I, T_{ex} is referred to as the "spin temperature," T_s. In LTE, the spin temperature is approximately the kinetic temperature of the gas. The Boltzmann factor gives us the level populations for a given excitation temperature.

The Boltzmann factor is:

$$\frac{n_{\rm u}}{n_\ell} = \frac{g_{\rm u}}{g_\ell} e^{-h\nu_{{\rm u}\ell}/kT_s} \,. \tag{2.7}$$

We know $\nu_{{\rm u}\ell}$ is at 21.11 cm/c, and $h\nu_{{\rm u}\ell}/k = 0.0682\,{\rm K}$. This low energy difference means that the electron is easily excited into the aligned state (by collisions or the CMB).

Let's talk about degeneracy. The electron spin can be aligned, or anti-aligned with the proton in the nucleus. The anti-parallel state has lower energy. For Hydrogen in the $L = 0$ state, the degeneracy is $2S + 1$. For the $S = 1$ state, it is 3, and for the $S = 0$ state it is 1. So...

$$\frac{n_{\rm u}}{n_\ell} = \frac{g_{\rm u}}{g_\ell} e^{-h\nu_{{\rm u}\ell}/kT_s} \simeq 3 \,. \tag{2.8}$$

What does this mean? The upper energy state has three times the atoms as the lower energy state, *and this is independent of the spin temperature!* Therefore, $n_u \simeq 3/4 n_T$ and $n_\ell \simeq 1/4 n_T$, where n_T is the total H I population in all states..

We can rewrite a general form for the emissivity

$$j_\nu = n_{\rm u} \frac{A_{{\rm u}\ell}}{4\pi} h\nu_{{\rm u}\ell} \phi_\nu \tag{2.9}$$

where ϕ_ν is again the normalized line profile, $\int \phi_\nu \, d\nu = 1$. We here are interested in ϕ_ν at line center where ϕ_ν is at a maximum. Plugging in our relations above:

$$j_\nu \simeq \frac{3}{16\pi} A_{{\rm u}\ell} h\nu_{{\rm u}\ell} n({\rm HI}) \phi_\nu \,. \tag{2.10}$$

Since most everything on the right hand side is a constant, the emissivity basically only depends on the density of H I atoms! This is a crazy result.

We can also rewrite our linear absorption coefficient:

$$\kappa_\nu = n_\ell \frac{g_{\rm u}}{g_\ell} \frac{A_{{\rm u}\ell}}{8\pi} \lambda_{{\rm u}\ell}^2 \phi_\nu \left[1 - e^{-h\nu_{{\rm u}\ell}/kT_s}\right] \,. \tag{2.11}$$

Since $e^{-h\nu_{{\rm u}\ell}/kT_s} \simeq 1$, stimulated emission is important! For example, for $\nu_{{\rm u}\ell} = 1420$ MHz and $T_s = 100$ K, we find $\exp(6.815 \times 10^{-4}) = 0.9993$. We can Taylor expand the exponential to get

$$\kappa_\nu \simeq \frac{3}{32\pi} A_{{\rm u}\ell} \frac{hc\lambda_{{\rm u}\ell}}{kT_s} n({\rm H\,I}) \phi_\nu \,. \tag{2.12}$$

The absorption coefficient $\kappa_\nu \propto 1/T_s$.

As the spin temperature goes up, the attenuation (and the optical depth) goes down. What this means is that cold H I does the absorbing. Draine derives the optical depth for Gaussian line profiles:

$$\tau_\nu = 2.190 \frac{N({\rm H\,I})}{10^{21}\,{\rm cm^{-2}}} \frac{100\,{\rm K}}{T_s} \frac{{\rm km\,s^{-1}}}{\sigma_V} e^{-u^2/2\sigma_V} \,, \tag{2.13}$$

where $N({\rm H\,I}) = \int n({\rm H\,I})\, ds$ is the column density and u is the velocity offset from line center, $u = v_r - v_{r,0}$.

We frequently have spin temperatures of 100 K and column densities $> 10^{21}\,\mathrm{cm}^{-2}$. Therefore, the attenuation and the optical depth can be very high. What is causing the optical depth? It is "self-absorption," absorption of H I by H I. This shows up as dark patches. [show papers].
http://adsabs.harvard.edu/cgi-bin/nph-bib_query?bibcode=2005ApJ...626..195G&db_key=AST
http://adsabs.harvard.edu/cgi-bin/nph-bib_query?bibcode=2002ApJ...566L..81J&db_key=AST

Imagine we have two clouds at different distances. Under what conditions will be observe self-absorption?

2.1 Kirchoff's Laws Applied to HI

H I can be observed in emission or in absorption. (This is true for many astrophysical emission mechanisms.) Kirchoff's laws tell us that we see emission from a diffuse gas and absorption from gas in front of a background continuum source. But there is *always* a background continuum source, the CMB! And we see lots of emission lines, so there must be some additional criterion when we get absorption lines.

In the radio regime, we often do on/off observations where the background off data are subtracted from the on-source data. This effectively removes the sky and telescope background.

Recall the solution to the equation of radiation temperature in the case of a constant source function in terms of brightness temperature:

$$T_B^{\mathrm{on}} = T_{B,0}e^{-\tau_\nu} + T_{\mathrm{s}}(1 - e^{-\tau_\nu})\,. \tag{2.14}$$

In many cases, the background $T_{B,0}$ is 2.7 K, from the CMB, but if there is a background radio source of course this is brighter than the CMB. I superscipted T_B with "on" to indicate that the observation is in the direction of interest. Observations always need a reference to remove sky and instrument artifacts.

Looking just off the H I cloud we only see the background unattenuated by the cloud:

$$T_B^{\mathrm{off}} \simeq T_{B,0}\,, \tag{2.15}$$

so the difference in the source and background radiation, is

$$\Delta T_B = T_{B,0}e^{-\tau_\nu} + T_{\mathrm{s}}(1 - e^{-\tau_\nu}) - T_{B,0} = (T_{\mathrm{s}} - T_{B,0})(1 - e^{-\tau_\nu}) \tag{2.16}$$

There are three regimes of spin temperature:
$T_{\mathrm{s}} = T_{B,0} : \Delta T_B = 0$. No line is visible.
$T_{\mathrm{s}} > T_{B,0} : \Delta T_B > 0$. Line appears in emission.
$T_{\mathrm{s}} < T_{B,0} : \Delta T_B < 0$. Line appears in absorption.

What this is telling us is that H I with low T_{s} can be seen in absorption against background continuum sources. This is the cold neutral medium (CNM). It is seen in emission in the absence of bright background sources. H I with high T_{s} is always seen in emission. This is

the "warm neutral medium," WNM. Observations along a sight line toward a bright quasar will therefore in general have both emission and absorption components.

The famous "Millennium" experiment explored H I absorption toward background radio sources: http://adsabs.harvard.edu/abs/2003ApJS..145..329H
http://adsabs.harvard.edu/abs/2003ApJ...586.1067H

2.2 Optically thin cloud - deriving the Column Density

If the optical depth is low and we can neglect absorption,

$$I_\nu(\tau_\nu) = I_{\nu,0} + \int j_\nu \, ds = I_{\nu,0} + \frac{3}{16\pi} A_{\mathrm{u}\ell} h\nu_{\mathrm{u}\ell} \phi_\nu N(\mathrm{H\,I}) \, . \tag{2.17}$$

If $I_{\nu,0}$ is known or can be estimated, we can integrate over the line profile to get the total H I emission.

$$\int [I_\nu(\tau_\nu) - I_{\nu,0}] \, d\nu = \frac{3}{16\pi} A_{\mathrm{u}\ell} h\nu_{\mathrm{u}\ell} N(\mathrm{H\,I}) \tag{2.18}$$

We can switch to antenna (T_A) or brightness (T_B) temperature (both are linear with I_ν) and evaluate the constants to get:

$$\int [T_B - T_{B,0}] \, dv = \Delta T_B \, dv = 54.89 \, \mathrm{K \, km \, s^{-1}} \frac{N(\mathrm{H\,I})}{10^{20} \mathrm{cm^{-2}}} \, . \tag{2.19}$$

The integration here is over the line profile, in velocity. This is a pretty incredible expression. What this is saying is that the integrated intensity (the intensity integrated over the H I line profile), is directly related to the H I column density. There is no dependence on spin temperature (assuming self-absorption is not important).

As a function of velocity, this is

$$\frac{dN(\mathrm{H\,I})}{dv} = 1.813 \frac{\Delta T_B(v)}{\mathrm{K}} \times \frac{10^{18} \mathrm{cm^{-2}}}{\mathrm{km \, s^{-1}}} \, , \tag{2.20}$$

assuming optically thin $\tau \lesssim 0.1$. Integrated over the line we have

$$N(\mathrm{H\,I}) = \frac{1.813 \times 10^{18} \, \mathrm{cm^{-2}}}{\mathrm{K \ km \, s^{-1}}} \int \Delta T_B(v) \, dv \, . \tag{2.21}$$

Since the column density is directly related to the mass, we can directly measure the H I mass of a Galaxy using observations of H I:

$$M_{\mathrm{HI}} = 4.945 \times 10^7 \ M_\odot \left(\frac{D_L}{\mathrm{Mpc}} \right)^2 \frac{F_{\mathrm{obs}}}{\mathrm{Jy \, MHz}} \, . \tag{2.22}$$

where D_L is the "luminosity distance" to the source (the distance derived from its magnitudes) and $F_{\rm obs}$ is the observed flux density in Jy. Note that, following Draine, I switched from temperature to flux density units. As you learned in radio astronomy (if you've taken it), one can easily switch back and forth between temperature and flux density. Instead of MHz, we can convert to velocity units using

$$\frac{\Delta\nu_{\rm obs}}{\nu_{\rm obs}} = \frac{\Delta v}{c}\frac{1}{1+z} \tag{2.23}$$

so

$$M_{\rm HI} = 2.343\times10^5\ M_\odot(1+z)^{-1}\left(\frac{D_L}{\rm Mpc}\right)^2\frac{\int F_{\rm obs}dv}{\rm Jy\ km\,s^{-1}}\,. \tag{2.24}$$

2.3 Background Radio source - Measuring T_s (Draine 8.3)

We can measure the spin temperature of a cloud if the background radiation source is bright. Assume we are observing two lines of sight through the same H I cloud: an "on" direction toward a background radio source and an "off" direction that just passes through the cloud. What makes a good background source? Something that emits a broad spectrum extending through 21 cm. A quasar is perfect. H II regions are great too. In such cases:

$$T_A^{\rm off} = T_{A,0}e^{-\tau_\nu} + T_s(1-e^{-\tau_\nu})\,. \tag{2.25}$$

Note that the "background" here is of actual emission; i.e., $T_A^{\rm off} \neq T_{A,0}$. Also note that I switched to units of antenna temperature T_A following Draine. On-source we have

$$T_A^{\rm on} = T_{\rm RS}e^{-\tau_\nu} + T_s(1-e^{-\tau_\nu})\,, \tag{2.26}$$

where RS stands for the (background) radio source and I have assumed no other background sources. so

$$T_A^{\rm on} - T_A^{\rm off} = \Delta T_A = (T_{\rm RS} - T_{A,0})e^{-\tau_\nu} \tag{2.27}$$

and we can directly estimate the optical depth if the absorption is strong. We similarly can get a relationship for the spin temperature.

With a little algebra (and switching from frequency to velocity), we find

$$\tau(v) = \ln\left[\frac{T_{\rm RS}-T_{\rm A,0}}{T_A^{\rm on}(v)-T_A^{\rm off}(v)}\right] \tag{2.28}$$

and

$$T_s = \frac{T_A^{\rm off}(v)T_{\rm RS} - T_A^{\rm on}(v)T_{\rm A,0}}{(T_{\rm RS}-T_{\rm A,0}) - (T_A^{\rm on}(v)-T_A^{\rm off}(v))} \tag{2.29}$$

When the absorption is strong, $(T_{\rm RS}-T_{\rm A,0})$ is larger than $(T_A^{\rm on}-T_A^{\rm off})$ and we can solve for both $T_s(v)$ and $\tau(v)$. If the absorption is weak, we can only get an upper bound on $\tau(v)$ and a lower bound on T_s.

Before, we derived the dependence of the column density on the line intensity, assuming $\tau_\nu \lesssim 0.1$. If this in not the case, we can correct for self-absorption if $\tau(v)$ and T_s are known:

$$\frac{dN(\mathrm{H\,I})}{dv} = 1.813 \frac{T_s \tau(v)}{\mathrm{K}} \times \frac{10^{18}\mathrm{cm}^{-2}}{\mathrm{km\,s}^{-1}}\,, \tag{2.30}$$

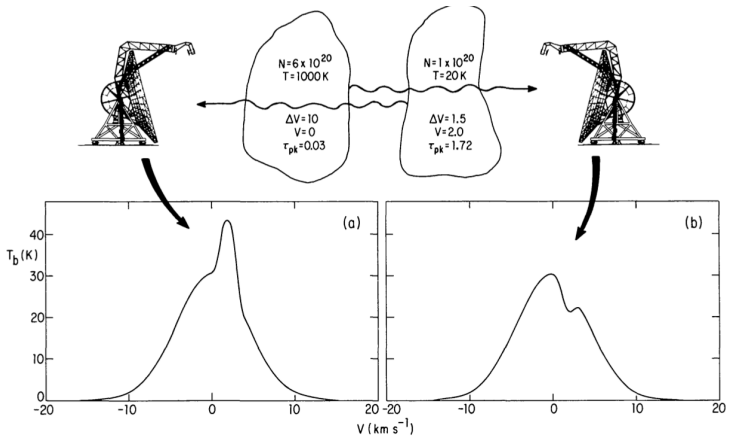


Figure 2.2: H I self-absorption schematic. The right panel shows self-absorption. (From Dicky & Lockman, 1990

2.4 Galactic HI

This is what a typical Galactic H I emission spectrum looks like when there is not much material along the line of sight:

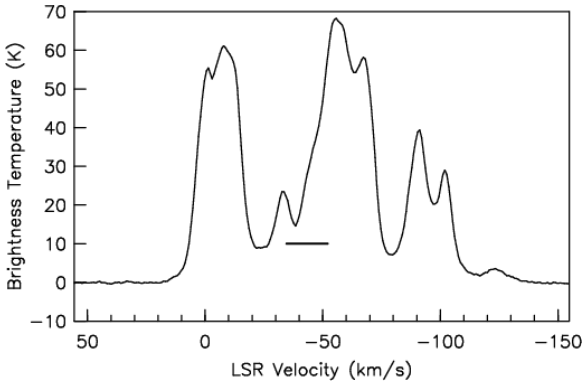


Figure 2.3: H I emission seen toward the outer Galaxy. The spectrum is taken in the "Local Standard of Rest," LSR frame. LSR is the circular velocity about the Galactic center at the location of the Sun.

Each Galactic H I cloud emits its own H I profile at a particular velocity. A typical Galactic H I spectrum does not show distinct lines. Why not? The main reason is that H I emits at all allowed velocities. Distinct H I lines are therefore smeared out by Galactic rotation. Along a line of sight, we get emission from all the various H I clouds, leading to the spectrum

in Figure 2.3. First we should understand the basics of Galactic rotation. This is just the basics, take a look at Carroll & Ostlie, Chapter 24 for more background.

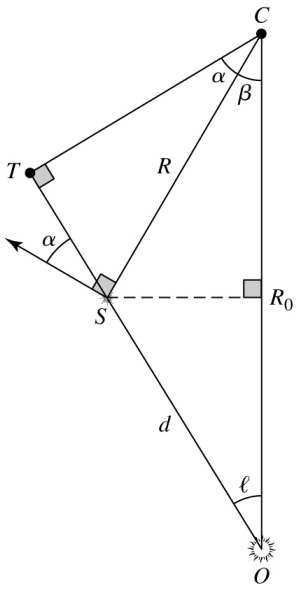


Figure 2.4: The geometry used when describing Galactic rotation.

Our Galaxy rotates. Defining the angular velocity curve:

$$\Omega(R) = \frac{\Theta(R)}{R}\,, \tag{2.31}$$

where $\Theta(R)$ is the velocity curve and R is the Galactocentric radius, the distance from the Galactic center. This of course assumes an axis-symmetric rotation curve. For circular rotation, we can derive

$$v_r = (\Omega(R) - \Omega_0) R_0 \sin\ell\,, \tag{2.32}$$

where R_0 is the Galactocentric radius of the Sun, 8.5 kpc[1], and ℓ is the Galactic longitude.

It turns our that our Galaxy rotates *differentially.* All parts of the disk in the midplane rotate with about the same linear speed, 235 $\mathrm{km\,s^{-1}}$ (Reid et al., 2014). As we look across the Galactic disk, we see parts of the Galaxy rotating radially toward and away from the Sun (with maximum, but never seen, values of 235 $\mathrm{km\,s^{-1}}$), and parts rotating completely tangentially at 0 $\mathrm{km\,s^{-1}}$. *Different measured radial velocities correspond to different distances.* Measuring the radial velocities of a cloud of gas for example leads to measurements of distances, assuming we know how the Galaxy rotates.

In order to make sense of the above equation, we need to know $\Omega(R)$. A good assumption backed up by observations is that $\Theta(R) \simeq const = 235$ $\mathrm{km\,s^{-1}}$, so $\Omega(R) = \frac{235\ \mathrm{km\,s^{-1}}}{R}$. Different forms of $\Omega(R)$ are called "rotation curves." Along a given line of sight, the law of cosines says

$$R^2 = R_0^2 + d^2 - 2R_0 d\cos\ell\,, \tag{2.33}$$

[1]This is the IAU value, although recent results have shown that 8.1 kpc is more accurate

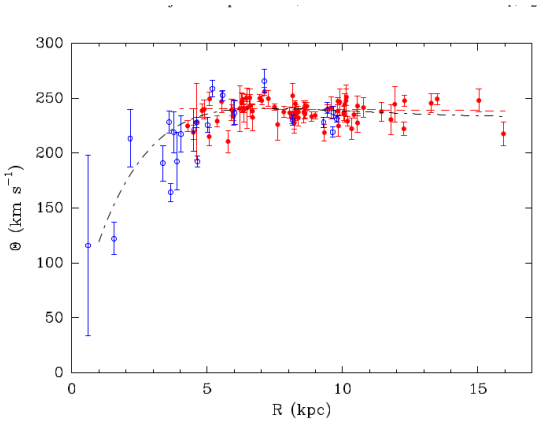


Figure 2.5: The rotation curve of the Milky Way as derived from massive star formation regions. Outside of a few kpc from the Galactic center, it is essentially flat. Figure from Reid et al. (2014).

so we can convert easily between d, the distance from the Sun, and R, the distance from the Galactic center. In general, for a given line of sight the absolute value of the radial velocity increases with decreasing R.

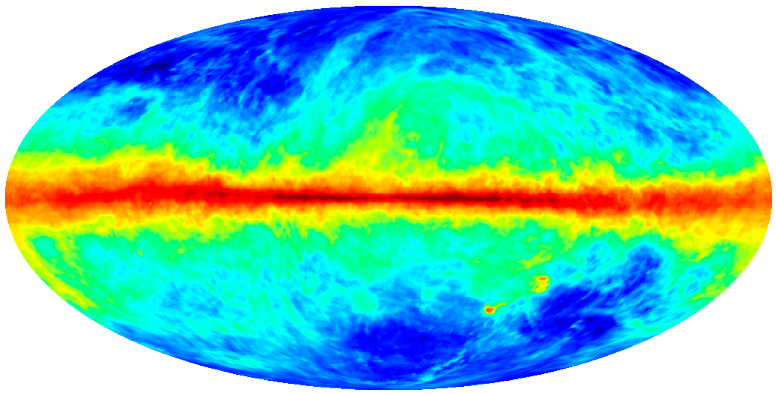


Figure 2.6: LAB H I integrated over the entire sky, in Galactic coordinates. The Galactic plane is clearly visible.

In the inner Galaxy, when the sight line crosses many values of R, the absolute value of the radial velocity versus distance curve increases, then decreases. Outside of the Solar orbit, the sign of the radial velocity curve is inverted. In the Outer Galaxy, as R increases with d, the absolute value of v_r just increases. Heliocentric distances are more difficult to determine in the inner Galaxy compared with the outer Galaxy, although Galactocentric distances can be found for any spectral line.

2.5 Extragalactic HI

Figure 2.9 shows a typical spectrum from an unresolved galaxy:

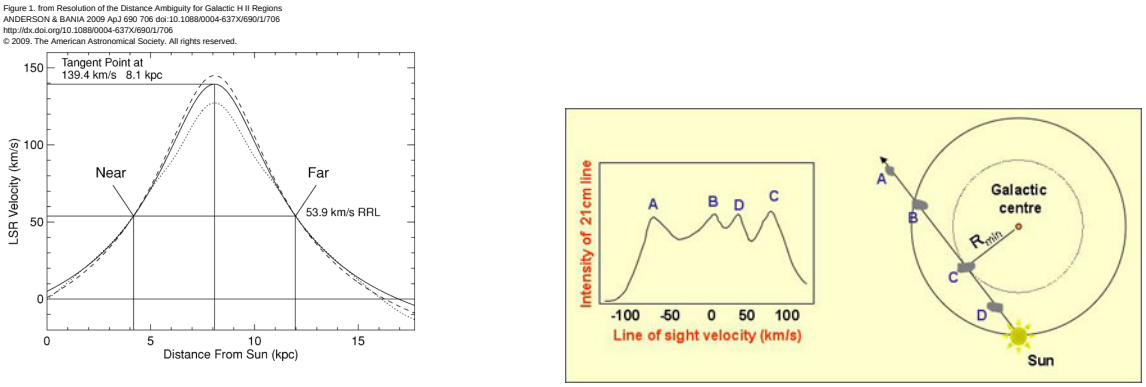


Figure 2.7: Inner Galaxy rotation curves.

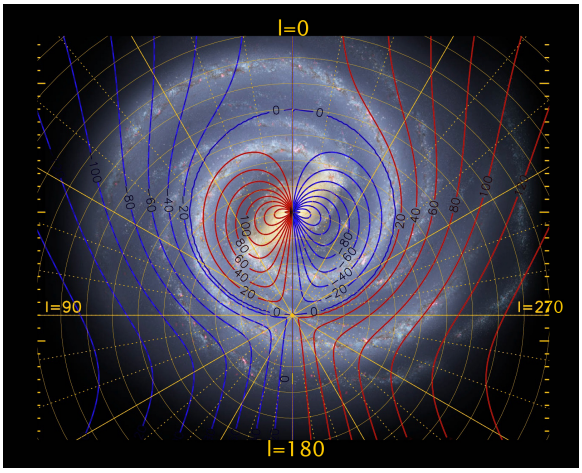


Figure 2.8: The velocity field of the Milky Way from our vantage point.

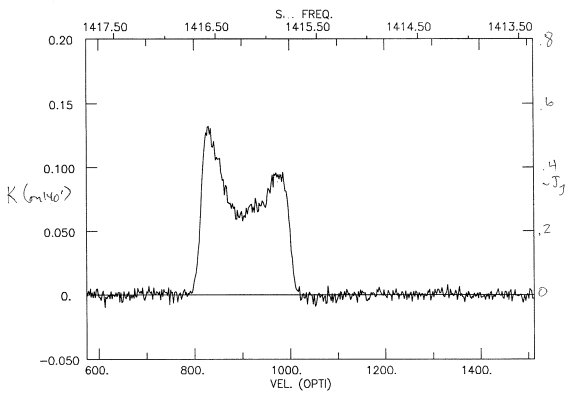


Figure 2.9: H I from an unresolved galaxy. The peaks ("horns") are caused by the low optical depth of H I and the flat rotation curve of galaxies.

Resolved observations of course have a wealth of information, but we will keep it simple here with an unresolved galaxy. In this case, all the emission from the galaxy is contained in this single spectrum. We get the "double-horn" profile because all spiral galaxies have, to first order, flat rotation curves. There is therefore lots of material moving at two particular velocities, the high and low velocities of the system. Your homework asks you about the optical depth of such a galaxy. An upcoming homework will ask you to relate this profile to the mass of the galaxy.

2.6 Draine Chapter 30: The Two-Phase HI

The temperature of H I (and all materials in the ISM) is determined by the balance between heating and cooling. As we saw in the Heiles & Trolund papers, there are two phases for H I: the CNM and WNM. These are in pressure equilibrium.

We can denote the "heating function Γ. What causes H I heating?
Ionization by cosmic rays
Photoionization of H and He by x-rays
Photoionization of dust grains by starlight U Photoionization of C, Mg, Si, Fe, etc by starlight UV
Heating by shock waves and other MHD phenomena
Draine gives expressions for all these.

Cooling on the other hand is dominated by metal lines, in particular, 63 μm OI and 158 μm CII. We can define a "cooling function" Λ in units of erg cm^{-3} s^{-1}; Figure 30.1 shows the cooling function as a function of temperature. You can see that the [OI] line dominates for higher densities and the [CII] line for lower densities. We'll get to these lines later.

The critical densities for these lines are $\sim 4 \times 10^3$ cm^{-3} for [CII] and $\sim 1 \times 10^5$ cm^{-3} for [OI]. [What does this mean? Collisional de-excition is unimportant] For these lines, the transitions are collisionally excited (more efficient as density increases), then spontaneously

emit. The heating and cooling balance leads to characteristic temperatures of the CNM and WNM.

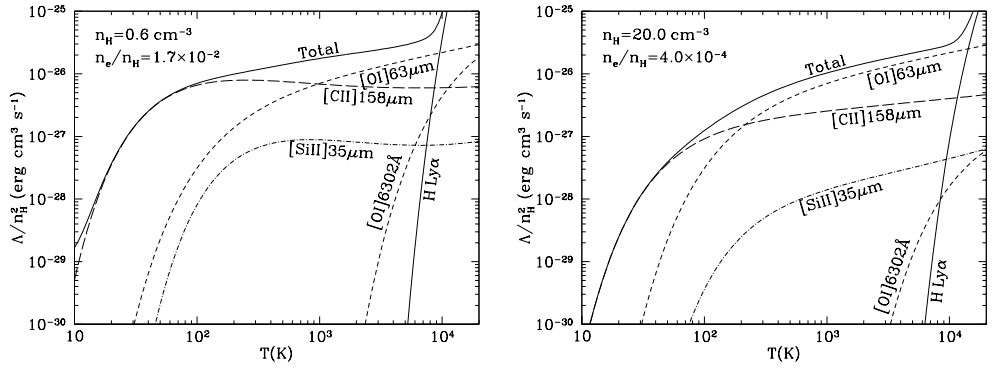


Figure 2.10: Cooling curves as a function of temperature for two densities in the ISM (from Draine book).

As shown in Draine Figure 30.2(a), we can calculate steady-state temperatures as a function of n_H. From this we see that the steady state temperature decreases as n_H increases. This is due to increases in the efficiency of cooling due to [OI] and [CII].

In a steady state ISM, everything must be in pressure equilibrium. Figure 31.2(b) shows that at low pressures, heating balances cooling at $T \simeq 6000\,\mathrm{K}$ (WNM) and at high pressures at $T \simeq 100\,\mathrm{K}$ (CNM). These two solutions are stable, but in between it is unstable.

It is often easier to think of density versus pressure, where the curve is functionally the same since $p/k = nT$. Looking at the unstable solution, if the density of the gas is somehow decreased, the pressure *increases* and the cloud expands. This expansion will lower the pressure until it gets to the WNM position. This is the "two-phase" H I.

It is entirely possible that the pressure of the ISM does not include the unstable range, but observationally we find that is more or less does. This does break down at high z off the plane though where essentially all the H I is in the WNM phase.

2.7 HI Oddities

We have by now explored most of the characteristics of H I in the ISM, but there are a number of peculiar features that you should know about. (see slides "HI_Oddities.pdf" for pretty pictures).

2.7.1 Superbubbles

https://en.wikipedia.org/wiki/Superbubble

Superbubbles are large H I shells filled with plasma in a high ionization state. They are created from the combined work of OB stars and supernovae.

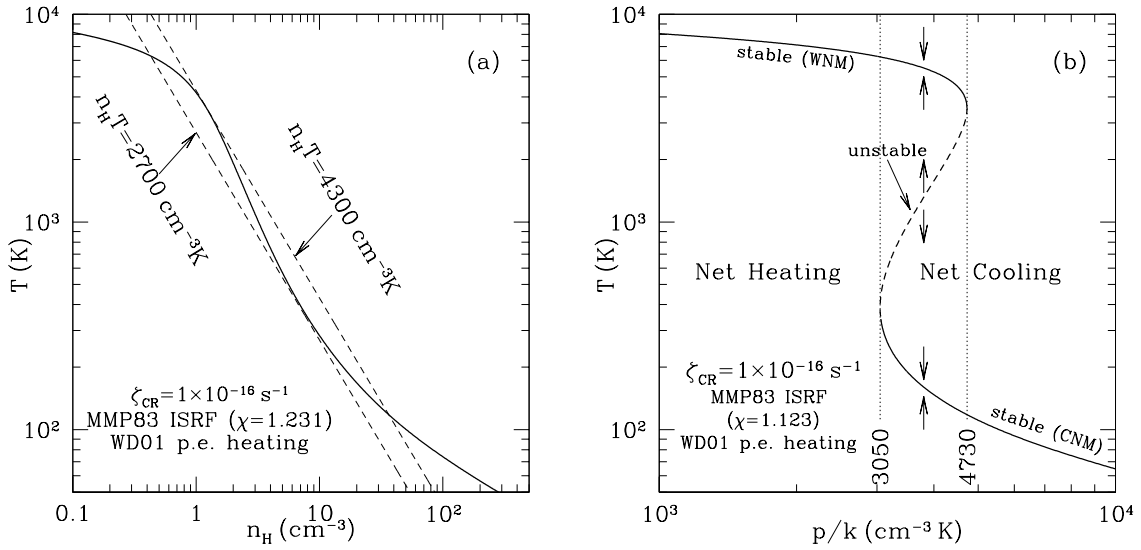


Figure 2.11: Heating by cosmic rays and photoelectric heating by dust grains produce the solid lines. For the given heating and cooling, we expect the neutral ISM gas to be found along these curves. Dashed and dotted lines show constant $n_H T$. In the right panel the same curves are shown. In the pressure range $3200 \lesssim p/k \lesssim 4400$ cm^{-3} K there are three possible equilibrium points: the CNM and WNM solutions, and also an unstable equilibrium solution.

Notable examples: Orion-Eradinus (see paper by Carl Heiles), Local Bubble

2.7.2 Riegel-Crutcher Cloud and Other Cold Absorbing Clouds

[NMG slides+paper, Gibson paper]
We can investigate the CNM in absorption if background continuum emission is strong enough. This usually limits our sight lines to those toward background AGN, BUT the Galactic center is a very strong continuum source as well. The Riegel-Cruther (R-C) Cloud is a cold HI cloud seen toward the Galactic center (see paper by Naomi McClure-Griffiths and the accompanying slides). We can work out the spin temperature of this cloud using the difference between two positions. First, we assume that the background continuum is behind the cloud, and uniform between on- and off- measurements. We allow for three HI components along the line of sight: foreground ($T_{s,fg}, \tau_{\nu,fg}$), the Reigel-Crutcher cloud (T_s, τ_ν), and background ($T_{s,bg}, \tau_{\nu,bg}$). The Galactic center continuum has temperature T_c.

At the location between the background cloud and the R-C cloud:

$$T_{\rm on,2} = T_c e^{-\tau_{\nu,bg}} + T_{s,bg}(1 - e^{-\tau_{\nu,bg}}) \tag{2.34}$$

Between the R-C cloud and the foreground cloud:

$$T_{\rm on,1} = T_{\rm on,2}e^{-\tau_\nu} + T_{\rm s}(1-e^{-\tau_\nu}) \tag{2.35}$$

$$T_{\rm on,1} = \left[T_c e^{-\tau_{\nu,bg}} + T_{s,bg}(1-e^{-\tau_{\nu,bg}})\right] e^{-\tau_\nu} + T_s(1-e^{-\tau_\nu}) \tag{2.36}$$

and finally at the observed location we have

$$T_{\rm on} = T_{\rm on,1}\left[e^{-\tau_{\nu,fg}} + T_{s,fg}(1-e^{-\tau_{\nu,fg}})\right] \tag{2.37}$$

$$T_{\rm on} = \left[T_c e^{-\tau_{\nu,bg}} + T_{s,bg}(1-e^{-\tau_{\nu,bg}})\right] e^{-\tau_\nu} + T_s(1-e^{-\tau_\nu})e^{-\tau_{\nu,fg}} + T_{\rm s,fg}(1-e^{-\tau_{\nu,fg}}) \tag{2.38}$$

$$T_{\rm on} = T_c e^{-(\tau_{\nu,bg}+\tau_\nu+\tau_{\nu,fg})} + T_{s,bg}(1-e^{-\tau_{\nu,bg}})e^{-(\tau_\nu+\tau_{\nu,fg})} + T_s(1-e^{-\tau_\nu})e^{-\tau_{\nu,fg}} + T_{s,fg}(1-e^{-\tau_{\nu,fg}}) \tag{2.39}$$

and similarly off the RC cloud

$$T_{\rm off} = T_c e^{-(\tau_{\nu,bg}+\tau_{\nu,fg})} + T_{s,bg}(1-e^{-\tau_{\nu,bg}})e^{-\tau_{\nu,fg}} + T_{s,fg}(1-e^{-\tau_{\nu,fg}}) \tag{2.40}$$

so therefore after some algebra

$$\Delta T = T_{\rm on} - T_{\rm off} = (T_s - T_{s,bg}(1-e^{-\tau_{\nu,bg}}) - T_c e^{-\tau_{\nu,bg}})e^{-\tau_{\nu,fg}}(1-e^{-\tau_\nu}) \tag{2.41}$$

Defining $p = T_{\rm s,bg}(1-e^{-\tau_{\nu,bg}})/T_{\rm off}$ and letting $\tau_{\nu,bg}$ and $\tau_{\nu,fg}$ be small relative to τ_ν we find

$$\Delta T = (T_s - T_c - pT_{\rm off})(1-e^{-\tau_\nu}) \tag{2.42}$$

and therefore

$$\tau_\nu = -\ln\left(1 - \frac{\Delta T}{T_s - T_c - pT_{\rm off}}\right) \tag{2.43}$$

This is all in Gibson et al. (2000). So we can get the optical depth with some knowledge or assumptions of the value of the p parameter.

2.7.3 High Velocity Clouds (incl. the Magellanic Stream)

High Velocity- clouds have velocities that deviate from that expected due to Galactic rotation. They are massive, reside in the halo, and take up a significant portion of the sky. https://en.wikipedia.org/wiki/High-velocity_cloud

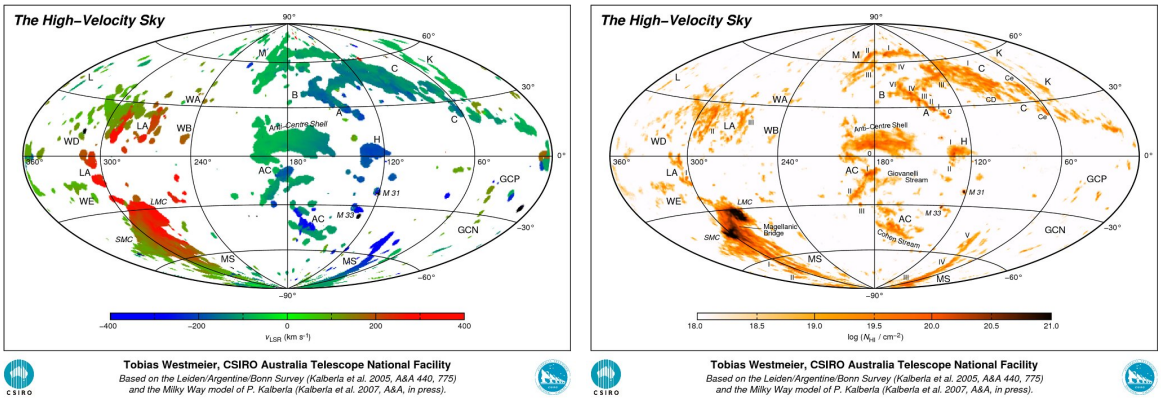


Figure 2.12: HVCs from http://www.atnf.csiro.au/people/Tobias.Westmeier/research_hvcsky.php made using LAB data.

Chapter 3

Molecules

There are over 100 molecules detected in the ISM. Each has strengths and weaknesses for learning about the ISM. [Rohlfs & Wilson Table]

How do molecules radiate? Molecular transitions are quantized, and therefore give rise to line emission. This line emission is quite a bit more complicated than that of H I because of the multitude of ways that molecules can radiate.

A non-exhaustive list of interactions within molecules includes:
1) Rotational lines due to rotation of a molecule about an axis
2) Vibrational lines due to the vibration of atoms in a molecule
3) "Inversion" transitions caused by QM tunneling within a molecule
4) Ro-vibrational (or "rotovibrational" = much more awesome name) transitions when the vibration frequency is the same as the rotation frequency. Each of these methods has associated energies, and therefore frequency ranges where they are detectable, as we will see later.

Let's just take a quick look at electronic transitions. Even in just the simplest case of diatomic hydrogen, we could have Coulomb electronic interactions of :
1) Electron 1 interacting with electron 2
2) Electron 1 interacting with proton 1
3) Electron 1 interacting with proton 2
4) Electron 2 interacting with proton 1
5) Electron 2 interacting with proton 2
6) Proton 1 interacting with proton 2. You can imagine how complicated things can get for larger molecules.

The Coulomb force dominates these interactions, but magnetic interactions occur as well. Furthermore, this list neglects the important vibrational and rotational emission mechanisms (described later). While for atomic hydrogen, we only have line emission from electronic transitions and the H I spin-flip transition, for molecular hydrogen, there are numerous additional

mechanisms for molecules that we have to consider.

It is worth noting that just like for atoms, molecular transitions have selection rules. Electric and magnetic dipole have $\Delta J = 0, \pm 1$. The others are given here:
https://en.wikipedia.org/wiki/Selection_rule.

3.1 Energies of Molecules

How can we make any progress in this situation? The key is that the energies of molecules can be separated into various terms. Also, we know that electrons are about 2000 times less massive than nuclei. We can therefore assume that the electrons are in motion, and the nuclei are fixed. This is valid because the electrons complete many "orbits" in the time it takes nuclei to complete only a small fraction of one. This "frozen nuclei" approximation is known as the "Born-Oppenheimer approximation." Born and Oppenheimer showed that molecular energies could be expressed as powers of $(m_e/M)^{1/2}$, where m_e is the electron mass and M is the nuclear mass.

Let's go through the most important molecular energies one-by-one. For all of these, I will assume that the length scale is approximately a_0, the Bohr radius, $a_0 = \hbar^2/(me^2)$. This is obviously incorrect, but allows us to get order-of-magnitude results.

Electronic Energies arise from the Coulomb interaction of the electrons to the nuclei, and have similar energies to those of atoms.

If we just use the energies from the Coulomb force:

$$E_{\text{electronic}} \simeq \frac{e^2}{a_0} \tag{3.1}$$

where e is the elementary charge and a_0 is the Bohr radius, the most probable distance between the proton and electron in a ground-state Hydrogen atom.

For reasons that will be apparent later, I am going to write this as

$$E_{\text{electronic}} \simeq \frac{e^2}{a_0}\left(\frac{m_e}{M}\right)^0 . \tag{3.2}$$

Vibrational Energies arise from oscillations of the nuclei around their equilibrium positions. This interaction is similar to that of two masses connected by a spring.

If we assume that the distance between the nuclei $r_0 \approx a_0$, then to change the separation between the nuclei by a distance a_0 we have to expend energy on the order of the electronic energy:

$$1/2 M\omega^2 a_0^2 \approx \frac{e^2}{a_0} . \tag{3.3}$$

Ignoring the factor of 2 and plugging in our expression for a_0,

$$\omega^2 \approx \frac{e^2}{M a_0^3}\,, \tag{3.4}$$

or

$$E_{\rm vibrational} = \hbar\omega \approx \frac{\hbar e}{M^{1/2} a_0^{3/2}} = \frac{\hbar e}{M^{1/2} a_0} \frac{m^{1/2} e}{\hbar} = \left(\frac{m_e}{M}\right)^{1/2} \frac{e^2}{a_0}\,. \tag{3.5}$$

Rotational Energies arise from motions of atoms about a rotation axis, the center of mass.

The moment of inertia for a diatomic molecule of size $r_0 \approx a_0$ is:

$$I \approx \mu r_0^2 \approx M a_0^2\,, \tag{3.6}$$

where μ is the reduced mass. The angular momentum L will be approximately $\hbar$. Therefore, using our expression for a_0

$$E_{\rm rotational} \approx \frac{L^2}{I} \approx \frac{\hbar^2}{M a_0^2} = \left(\frac{m_e}{M}\right) \frac{e^2}{a_0} \tag{3.7}$$

Since, $m/M \approx 10^{-4}$,

$$\frac{E_{\rm electronic}}{E_{\rm vibrational}} \approx 10^2 \tag{3.8}$$

and

$$\frac{E_{\rm vibrational}}{E_{\rm rotational}} \approx 10^2\,. \tag{3.9}$$

As a result, electronic transitions are found in the optical/UV, vibrational in the IR, and rotational in the radio (mm for light molecules and cm for heavier ones).

For the rest of our discussion on molecules, we will mostly ignore electronic transitions.

3.1.1 Rotational Transitions [From Cormac Purcell's Thesis]

Classically, rotation is described in terms of the moment of inertia I about a particular axis:

$$I = \sum m_i\, r_i^2, \tag{3.10}$$

where r_i is the perpendicular distance of the $i^{\rm th}$ atom from the axis of rotation, and m_i is its mass. A molecule has associated moment of intertias about three axes (in three dimensions), I_a, I_b, and I_c. The total kinetic energy is therefore:

$$E = \frac{1}{2}\left[\, I_a\, \omega_a^2 + I_b\, \omega_b^2 + I_c\, \omega_c^2 \,\right], \tag{3.11}$$

where $\omega_{\rm a}$ is the angular velocity in radians/sec about an axis 'a'. Equation 3.11 may be written in terms of the classical angular momentum, $L_a = I_a\, \omega_a$:

$$E = \frac{L_a^2}{2I_a} + \frac{L_b^2}{2I_b} + \frac{L_c^2}{2I_c}, \tag{3.12}$$

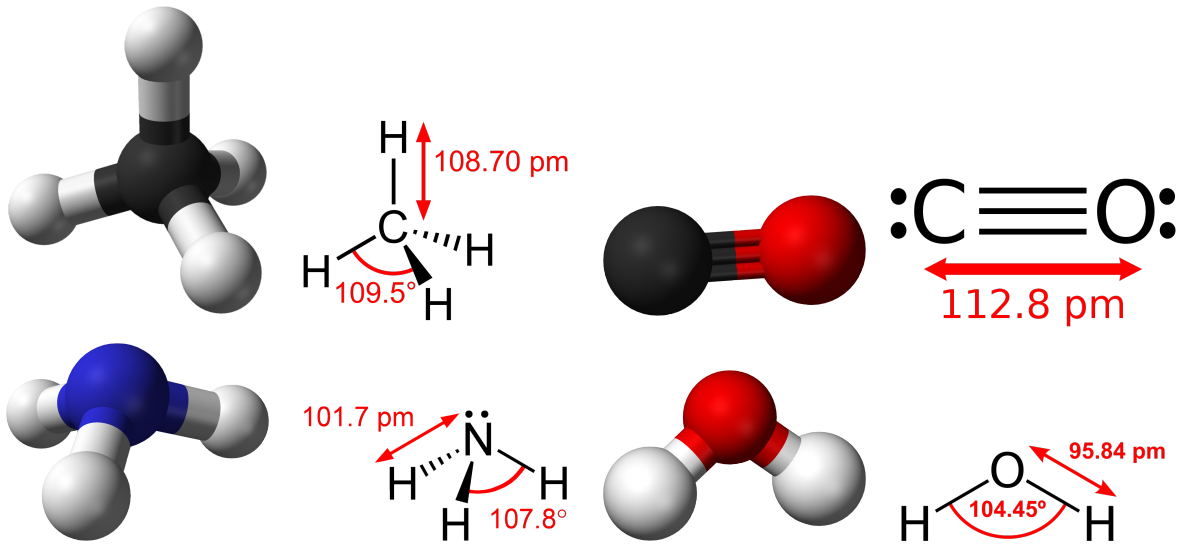


Figure 3.1: Diagrams for various molecules (from Wikipedia). Top left shows methane (CH_4; symmetric rotor; top right shows carbon monoxide (CO; linear); bottom left shows ammonia (NH_3, symmetric rotor); bottom right shows water (H_2O, asymmetric rotor).

with the magnitude of the total angular momentum given by $L^2 = L_a^2 + L_b^2 + L_c^2$.

We can classify molecules based on their symmetry, because as we will see this symmetry dictates the form of the rotational transitions.

Spherical Rotors:	$I_a = I_b = I_c$,	e.g.: CH_4, SiH_4.
Linear Rotors:	$I_a = 0, I_b = I_c$,	e.g.: CO, HCO+, HCN, HNC, N_2H^+
Symmetric Rotors:	$I_a = I_b \neq I_c$,	e.g.: NH_3, CH_3CN, CH_3Cl.
Asymmetric Rotors:	$I_a \neq I_b \neq I_c$,	e.g.: H_2O, CH_3OH.

To emit or absorb radiation efficiently the molecule must have a permanent dipole moment, μ[1]. Usually the dipole is an electric moment due to the asymmetric distribution of positive and negative charges on a molecule. The electronic charge on spherical and homo-nuclear linear molecules is evenly distributed and these molecules do not exhibit rotational dipole transitions. Instead much weaker quadrupole lines are observed in the infrared, due to simple collisional excitation. The quadrupole arises from the interaction of external electromagnetic fields with the slightly asymmetric charge distribution in the nucleus.

3.1.2 Spectroscopic Notation

Conveniently, molecular spectroscopic notation is similar to that for atoms:

$$^{(2\Sigma+1)}\Lambda_{J,u/g}\,, \tag{3.13}$$

[1]Molecules without a permanent dipole moment can couple to external radiation fields via higher moments, e.g., H_2 and N_2 exhibit quadrupole transitions, however these lines are usually weak.

$\Lambda = \Sigma, \Pi, \Delta...$, for $\Lambda = 0, 1, 2...$, where $\Lambda\hbar$ is the projection of the electron orbital angular momentum onto the inter-nuclear axis (like $\mathcal{L}$ for atoms), $\Sigma\hbar$ is the projection of the electron spin angular momentum onto the inter-nuclear axis (like S for atoms), J is the projection of the total angular momentum along the inter-nuclear axis (sometimes denoted Ω), and u/g gives the parity.

Just like in atoms, different states of Λ have different names. These follow the same letters as for atoms, but using the capital Greek alphabet. For example, $\Lambda = 0, 1, 2$, and 3 give rise to molecular electronic states Σ, Π, Δ, and Φ.

As you may expect, the quantum number $\vec{J} = \vec{\Lambda} + \vec{\Sigma}$. Molecules with even values for J are called "ortho," and those with odd values are called "para." Otho and para molecules behave as distinct species, with distinct energy level diagrams. This is most important for H_2O, where the frequency differences of ortho and para transitions are important.

While not included in the notation, K is the projection of the total angular momentum onto the principal axis; $K = 0$, ±1, ±2, $\ldots \pm J$.

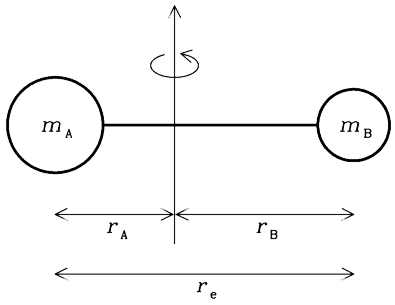
3.2 Linear Molecules: CO Rotational Transitions

There are a wealth of different molecules, each with its own Einstein A coefficients and critical densities, and of course they tell us different things about the ISM. We are going to focus initially on CO, which is the most often used molecular gas tracer, due to its high abundance and low critical density. We will first talk about rotational transitions, because they are conceptually the easiest.

Why are we using CO? Hydrogen is the most abundant element, and when it is in the form of *atomic* hydrogen, it is relatively easy to observe. However, at the high densities where stars form, hydrogen tends to be molecular rather than atomic, and H_2 is extremely hard to observe directly. To understand why, we must review the quantum structure of H_2.

[From Pogge's notes] While the dominant molecular species in the ISM is H_2, because it is a homo-nuclear linear molecule with no permanent dipole moment all of the low-lying energy levels are quadrupole transitions with small transition probabilities (A-values) and relatively high excitation energies. Because it is homonuclear, there is no permanent dipole moment and therefore no rotational transitions, only vibrational. The high excitation energies mean that these transitions are only excited at high temperatures uncommon for most of the ISM or in strong UV radiation fields (i.e., fluorescence). Thus the most abundant molecule in the ISM, carrying most of the mass and playing a key role in excitation, thermal balance, and gas-phase chemistry, is virtually invisible to direct observation.

The conclusion of this analysis is that, for typical conditions in star-forming clouds, we cannot observe the most abundant species, H_2, in emission. Instead, we are forced to observe proxies. We will talk more about H_2 at the end of our work on molecules.



In the sections below, we are going to use CO to explain some ways that we can use molecular line observations to derive physical quantities.

3.2.1 Rotational Energies

Consider a rigid diatomic molecule whose two atoms have masses m_A and m_B and whose centers are separated by the equilibrium distance r_e. The individual atomic distances r_A and r_B from the center of mass must obey

$$r_e = r_A + r_B \quad \text{and} \quad r_A m_A = r_B m_B \,. \tag{3.14}$$

The rotational energy is

$$E_{rot} = \frac{1}{2} I \omega^2 \,, \tag{3.15}$$

or in terms of the orbital angular momentum $L = I\omega$,

$$E_{rot} = \frac{1}{2} \frac{L^2}{I} \,. \tag{3.16}$$

The moment of inertia is simply

$$I = \sum m_i r_i^2 = (m_A r_A^2 + m_b r_B^2) \,. \tag{3.17}$$

So the net result is that we can derive a relationship between the rotational energy and the orbital frequency given the masses and separations of the atoms. Note that the energies are necessarily quantized because L is quantized.

Just like for atoms, molecules have selection rules. Here, selection rules require that $\Delta J = \pm 1$. CO rotational energies are often given as:

$$E_J = hBJ(J+1) \,, \tag{3.18}$$

with the "rotational constant" $B = \frac{h}{8\pi^2 I} = 57.63596$ GHz, where I is the moment of inertia. This is known as a "rigid rotor" because as it rotates faster and faster the distance between C and O is preserved. Rigid rotors therefore give rise to rotational energies

$$\Delta E = hBJ(J+1) - hB(J-1)J = hB(J^2 + J - J^2 + J) = 2hBJ \tag{3.19}$$

or frequencies

$$\nu_{ij} = 2BJ\,, \tag{3.20}$$

where J is the upper rotational level.

Note that the frequencies for rigid rotor transitions are simply linear with J! We therefore get a "ladder" of emission.

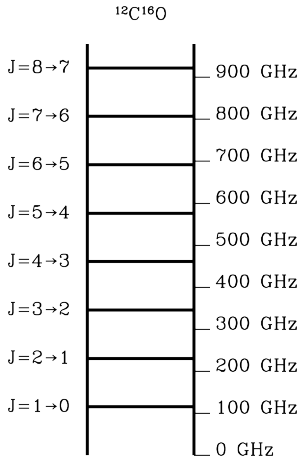


Figure 3.2: The CO "ladder" (from ERA course).

The rotation of CO causes centrifugal distortion. At higher rotational energies we can think of the molecule as being pulled apart. This raises the moment of inertia and therefore lowers the energy. To account for this, centrifugal distortion correction terms are added to the rotational energy levels of the diatomic molecule.

$$E_J = h[BJ(J+1) - DJ^2(J+1)^2] \tag{3.21}$$

and therefore the lines are at

$$\nu \simeq 2BJ(J+1) - 4DJ^2(J+1)^2 \tag{3.22}$$

For CO the value of D is 0.1846066 MHz. The relative value of this constant tells you roughly about its importance relative to B (which is about 50 GHz). An assumption underlying these expressions is that the molecular vibration follows simple harmonic motion. In the harmonic approximation the centrifugal constant D can be derived as

$$D = \frac{h^3}{32\pi^4 I^2 r^2 kc} = 0.1835055 \text{ MHz}\,, \tag{3.23}$$

which is close to the true value above.

In real spectroscopy, one simply looks up the frequency of the transition of interest, which has been derived in a lab.

3.2.2 Is CO Optically Thick?

[Draine Chapter 19.3] In order to use our observations to derive physical parameters, we must determine if CO is optically thin or thick. We know that the linear absorption coefficient κ_ν is

$$\kappa_\nu = n_\ell \frac{g_\mathrm{u}}{g_\ell} \frac{A_{\mathrm{u}\ell}}{8\pi} \lambda_{\mathrm{u}\ell}^2 \phi_\nu \left(1 - \frac{n_\mathrm{u}}{n_\ell}\frac{g_\ell}{g_\mathrm{u}}\right) \tag{3.24}$$

(Draine Eqn. 8.4). The $\frac{n_\mathrm{u}}{n_\ell}\frac{g_\ell}{g_\mathrm{u}}$ term is sometimes written as $e^{-h\nu/kT_\mathrm{ex}}$. Remember: as written, the "1" is from absorption and the "$\frac{n_\mathrm{u}}{n_\ell}\frac{g_\ell}{g_\mathrm{u}}$" is from stimulated emission. Assuming Gaussian lines,

$$\phi_\nu = \frac{1}{2\pi}\frac{c}{\nu_{\mathrm{u}\ell}}\frac{1}{\sigma_V} e^{-u^2/2\sigma_V^2} \tag{3.25}$$

so

$$\kappa_\nu = n_\ell \left(1 - \frac{n_\mathrm{u}}{n_\ell}\frac{g_\ell}{g_\mathrm{u}}\right) \frac{\lambda^2}{8\pi}\frac{g_\mathrm{u}}{g_\ell} A_{\mathrm{u}\ell} \frac{1}{\sqrt{2\pi}}\frac{1}{\sigma_V} e^{-u^2/2\sigma_V^2} \tag{3.26}$$

Since $\tau_0 = \int \kappa_0\, ds$, where the 0 subscript specifies line center,

$$\tau_0 \simeq \kappa_{\nu_0}\, R = n_\ell\, R \left(1 - \frac{n_\mathrm{u}}{n_\ell}\frac{g_\ell}{g_\mathrm{u}}\right) \frac{\lambda^2}{8\pi^{3/2} b}\frac{g_\mathrm{u}}{g_\ell} A_{\mathrm{u}\ell}\,, \tag{3.27}$$

where R is the cloud diameter (Draine says radius in my version of the text, but I think diameter is correct), and $b^2 \equiv 2\sigma_V^2$. In typical units, after plugging in $A_{10} = 7.16 \times 10^{-8}\,\mathrm{s}^{-1}$, this reduces to

$$\tau_0 = 297\, n_3\, R_{19} \left[\frac{n(\mathrm{CO})/n_H}{7\times10^{-5}}\right] \left[\frac{n(J=0)}{n(\mathrm{CO})}\right] \left(\frac{2\ \mathrm{km\,s^{-1}}}{b}\right) \left(1 - \frac{n_\mathrm{u}}{n_\ell}\frac{g_\ell}{g_\mathrm{u}}\right)\,, \tag{3.28}$$

where n_3 and R_{19} are the density in units of 10^4 cm^{-3} and diameter in units of 10^{19} cm. But to use this we need an estimate of the fraction of CO in the $J = 0$ state. You all know where this is going....

We can modify the Boltzmann factor for CO using the partition function:

$$\frac{n(\mathrm{CO}, J)}{n(\mathrm{CO})} = \frac{(2J+1)e^{-hBJ(J+1)/kT_\mathrm{ex}}}{\Sigma_J (2J+1) e^{-hBJ(J+1)/kT_\mathrm{ex}}} \tag{3.29}$$

T_ex here is sometimes called the "rotation temperature," in the same way that we use "spin temperature" for H I. We can approximate

$$Z = \sum_J (2J+1) e^{-hBJ(J+1)/kT_\mathrm{ex}} \simeq \left[1 + (kT_\mathrm{ex}/hB)^2\right]^{1/2} \tag{3.30}$$

We know $B/k = 2.77\,\mathrm{K}$ for ^{12}CO, and since $\left(1 - \frac{n_\mathrm{u}}{n_\ell}\frac{g_\ell}{g_\mathrm{u}}\right) = (1 - e^{-hBJ(J+1)/kT_\mathrm{ex}})$, therefore

$$\tau_0 \simeq 297\, n_3\, R_{19} \left[\frac{n(\mathrm{CO})/n_H}{7\times10^{-5}}\right] \left[\frac{(1 - e^{-5.53\,\mathrm{K}/T_\mathrm{ex}})}{[1 + (T_\mathrm{ex}/2.77)^2]^{1/2}}\right] \left(\frac{2\ \mathrm{km\,s^{-1}}}{b}\right)\,. \tag{3.31}$$

or if we can assume $T_{\rm ex} = 8\,{\rm K}$,

$$\tau_0 \simeq 46\, n_3\, R_{19} \left[\frac{n({\rm CO})/n_H}{7\times10^{-5}}\right]\left(\frac{2\ {\rm km\,s^{-1}}}{b}\right)\,. \tag{3.32}$$

The punchline here is that ^{12}CO is usually optically thick, for reasonable values of n, R, and b.

3.2.3 Critical Density for CO and Radiative Trapping

As we just saw, CO is frequently optically thick. The radiation cannot escape from the cloud, and is trapped. Therefore, emission from the upper level is absorbed within the cloud, leading to a strong stimulated emission term. This emission is therefore observed to be less intense, and the level populations are skewed such the the upper level is more populated. The excitation temperature is *above* the kinetic temperature of the gas. This is known as **radiative trapping** (Draine Ch. 19).

Draine works through a nice derivation called the **escape probability approximation**. A full treatment would be very involved, but it turns out that the approximation is quite good. The result is that the average escape probability $\langle B\rangle_{\rm cloud}$ can be approximated as

$$\langle B\rangle_{\rm cloud} = \frac{1}{1+0.5\tau_0} \tag{3.33}$$

As you can imagine, this makes our old definition of the critical density useless. In the case of optically thick clouds, we have to modify the critical density:

$$n_{\rm crit} = \frac{\langle\beta_{{\rm u}\ell}\rangle A_{{\rm u}\ell}}{C_{{\rm u}\ell}}\,, \tag{3.34}$$

where β is the escape probability. Authors have come up with an analytical expression for the collisional de-excitation parameter

$$C_{{\rm u}\ell} \simeq 6\times10^{-11}\,T_2^{0.2}\,{\rm cm^3\,s^{-1}} \tag{3.35}$$

for 10 K $\lesssim T \lesssim$ 250 K (Flower & Launay 1985; Flower 2001). For a cloud with density 1000 cm^{-3}, $R_{19} = 1\,{\rm cm}$, $b = 2\ {\rm km\,s^{-1}}$ (again, $b^2 \equiv 2\sigma^2$), and $T_{\rm ex} = 8\,K$, $\tau_0 \simeq 50$ and $\langle\beta\rangle = 1/(1+25) \simeq 0.04$. Therefore,

$$n_{{\rm crit},{\rm H_2}}({\rm CO}, J=1) = \frac{\langle\beta_{{\rm u}\ell}\rangle A_{{\rm u}\ell}}{C_{{\rm u}\ell}} \simeq 50 T_2^{-0.2}\ {\rm cm^{-3}}. \tag{3.36}$$

This shows that the $J = 1$ level of CO is "thermalized" in molecular clouds with $n_{H_2} \gtrsim 10^2$ cm^{-3}. Thermalized means that $T_{\rm ex} \simeq T_K$. You can get a *larger* excitation temperature at a *lower* density, effectively lowering the critical density. For optically thin clouds, we have $n_{{\rm crit},{\rm H_2}}({\rm CO}, J=1) = 1100 T_2^{-0.2}$ and we require much higher critical densities.

(This discussion is frequently absent in texts, but is very important because it shows that CO is thermalized at densities far below the critical density. For further reading, see the Shirley and Magnum articles in the "Articles" folder; also see below for a couple results from the Shirley paper.)

3.2.4 CO Isotopologues

Since ^{12}CO is optically thick, it cannot tell us all we need about the ISM. Luckily, there are a few isotopologues (same atoms, different atomic weights) of CO that are frequently used: ^{13}CO, C^{18}O, and C^{17}O.

For rotational transitions,

$$\nu = \frac{hJ}{4\pi^2 \mu r_e^2} \tag{3.37}$$

We can see that as the reduced mass μ increases, the frequency decreases. The ^{12}CO line is at 115.271208 GHz. Assuming r_e does not change, ^{13}CO line must therefore be at [from NRAO course]

$$\frac{\mu(^{13}\mathrm{C}^{16}\mathrm{O})}{\mu(^{12}\mathrm{C}^{16}\mathrm{O})} = \frac{13 \times 16/(13+16)}{12 \times 16/(12+16)} = 1.0460 \tag{3.38}$$

so

$$\nu_{10}(^{13}\mathrm{C}^{16}\mathrm{O}) = \nu_{10}(^{12}\mathrm{C}^{16}\mathrm{O}) \times \left[\frac{m(^{13}\mathrm{C}^{16}\mathrm{O})}{m(^{12}\mathrm{C}^{16}\mathrm{O})}\right]^{-1} = 115.271208/1.0460 = 110.204 \text{ GHz} \tag{3.39}$$

This is not quite correct though because r_e of course does also change (see centrifugal distortion term from before). The actual ^{13}CO $J = 1 \rightarrow 0$ frequency is 110.201354 GHz. The ratio of ^{12}CO to ^{13}CO should be around 90 [from Pogge notes] and 45 from Langer & Penzias (1990) (this is a mean value for Galactocentric radii between 4 and 8 kpc). For reference, C^{18}O is at 109.782156 GHz and C^{17}O is at 112.359275 GHz.

We want to have a nice, optically thin tracer to use to derive cloud masses, etc. ^{13}CO is ok for this, although it turns out to have an optical depth near unity, so sometimes C^{18}O is better. We can assume is that ^{13}CO and ^{12}CO both arise in the same regions, and so share the same excitation temperature. If there are significant chemical fractionation effects, this assumption could be invalid (e.g., if local chemistry affects the creation/destruction of ^{13}CO differently than that of ^{12}CO).

3.2.5 Other Linear Molecules

The molecules most like CO have their atoms aligned along one axis. There are many other linear molecules, notably all diatomic molecules (CS, CF, HF, SiO, etc.). Other good linear molecules are the carbon chain molecules like HCN, $HC_{11}N$, etc.

Just like CO, we get frequency ladders for all of these. It is important to note that as the mass goes up, the moment of inertia increases and the energies decrease. Most of these rotational transitions are in the radio regime, but the more massive molecules are found at lower frequencies. One example is the $J = 1 \rightarrow 0$ transition of CS, which is found at 43 GHz, compared to the same transition for CO at 115 GHz.

Each molecule will have its own critical density. Radiative trapping is most important for CO, but is also important for many other molecules.

3.2.6 A few more notes on $n_{\rm crit}$

Shirley (2015) did a nice analysis of the critical density. He did the full multi-level solutions, which is a big improvement over the typical treatments.

I found a few important nuggets from this text (quoted here):

- As the frequency of the transition increases, the critical density moves farther up the $T_{\rm ex}$ curve.
- At optical wavelengths (the high frequency limit), critical density is interpreted as the density at which an atomic forbidden line is quenched by collisions meaning that a collisional de-excitation occurs before a photon can be generated from spontaneous emission. In the optical, quenching occurs when $T_{\rm ex}$ is thermalized ($T_{\rm ex} \rightarrow T_K$). At radio wavelengths (the low frequency limit) the critical density occurs at different positions along the sigmoid $T_{\rm ex}$ curve for different molecular tracers and different transitions. ... neither the radio interpretation of critical density being the density at which an emission line is excited nor the optical interpretation of critical density being the density at which an emission line is quenched is appropriate.
- $n_{\rm crit}$ provides only a rough estimate of the densities traced and should not be over-interpreted. More sophisticated tools such as the Contribution Function (Tafalla et al. 2006) may be used to determine the various contributions to the observed line profile along the line-of-sight (see Pavlyuchenkov et al. [2008] for a detailed analysis of these techniques). Ultimately, if one wants to understand their observed spectra, radiative transfer modeling with publicly available codes are a fast and efficient way to determine the physical properties of a region.

3.3 Using CO to Derive Physical Parameters

Below, I show how we can use CO line observations to derive physical properties of molecular clouds. The same principles apply to many molecular lines.

3.3.1 $T_{\rm ex}$ and T_k

Since we have shown that ^{12}CO is usually optically thick (Equation 3.32), the excitation temperature is easily derived. We know very approximately that for optically thick lines $\Delta T_B \simeq T_{\rm ex} \simeq T_k$. For CO, however, the Rayleigh-Jeans approximation is not entirely valid ($\nu \simeq 100$ GHz for $J = 1 - 0$), and we should do the complete solution.

At 100 GHz, the R-J limit doesn't quite apply. We therefore have expressions that are considerably uglier. Let's revisit our definition of the brightness temperature:

$$T_B \equiv \frac{c^2}{2k\nu^2} F_\nu \,, \tag{3.40}$$

where F_ν is the flux density (usually in Jy beam^{-1}). Therefore, with this definition, the

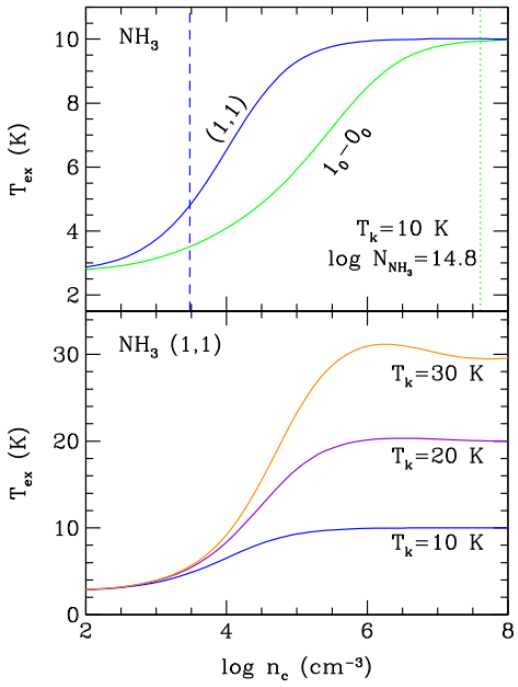


Figure 3.3: The excitation temperature for NH_3 as a function of density of particles with which it collides (n_c). Vertical lines show the critical density. From Shirley (2015).

solution to the equation of radiative transfer in the case of a grey-body background leads to

$$I_\nu = B_\nu(T_{BG})e^{-\tau_\nu} + B_\nu(T_{\rm ex})\left(1 - e^{-\tau_\nu}\right) \tag{3.41}$$

or, using brightness temperature,

$$T_B = \frac{h\nu}{k}\left[\frac{1}{e^{h\nu/kT_{BG}} - 1}e^{-\tau_\nu} + \frac{1}{e^{h\nu/kT_{\rm ex}} - 1}\left(1 - e^{-\tau_\nu}\right)\right] \tag{3.42}$$

If we take an off-source direction,

$$T_B^{\rm off} = \frac{h\nu}{k}\frac{1}{e^{h\nu/kT_{BG}} - 1} \tag{3.43}$$

[Rohlfs & Wilson] So therefore:

$$\Delta T_B = (1 - e^{-\tau_\nu})\frac{h\nu}{k}\left[\frac{1}{e^{h\nu/kT_{\rm ex}} - 1} - \frac{1}{e^{h\nu/kT_{BG}} - 1}\right] \tag{3.44}$$

when the background radiation is assumed to be from the CMB $T_{BG} = 2.7\,\mathrm{K}$, $h\nu/k = hBJ(J{+}1) = 5.5\,\mathrm{K}$ for the optically thick $J = 1 \rightarrow 0$ line, and if we assume LTE ($T_{\rm ex} = T_K$):

$$T_{\rm ex} = \frac{5.5}{\ln\left(1 + 5.5/(\Delta T_B(^{12}\mathrm{CO}) + 0.82)\right)}\,. \tag{3.45}$$

If $\Delta T_B = 10$ K, $T_{\rm ex} = 13.4$ K.

3.3.2 An Alternate way of Deriving $T_{\rm ex}$ [Kwok 9.2]

The easiest way to derive $T_{\rm ex}$ is to assume that ^{12}CO is optically thick, and to use Equation 3.45. This may be ok in some cases, but the assumption of optically thick CO does break down at high J, and furthermore may not be a great assumption for more diffuse clouds. We can, however, use line ratios of various transitions for a single isotopologue to derive the excitation temperature.

We have

$$\frac{n({\rm CO},J)}{n({\rm CO})} = \frac{(2J+1)e^{-E_J/kT_{\rm ex}}}{Z(T_{\rm ex})}\,. \tag{3.46}$$

If we assume that all lines are optically thin, $n({\rm CO},J) \propto N({\rm CO},J)$, then

$$\frac{N_J}{N_T} = \frac{(2J+1)e^{-E_J/kT_{\rm ex}}}{Z(T_{\rm ex})}\,. \tag{3.47}$$

$$\ln N_J - \ln N_T = \ln(2J+1) - \frac{E_J}{kT_{\rm ex}} - \ln(Z(T_{\rm ex}))\,. \tag{3.48}$$

$$\ln N_J - \ln(2J+1) = \ln N_T - \ln(Z(T_{\rm ex})) - \frac{E_J}{kT_{\rm ex}}\,. \tag{3.49}$$

$$\ln\left(\frac{N_J}{2J+1}\right) = \ln\left(\frac{N_T}{Z(T_{\rm ex})}\right) - \frac{E_J}{kT_{\rm ex}}\,. \tag{3.50}$$

$$\tag{3.51}$$

Therefore, we have a linear equation! If we plot $\ln\left(\frac{N_J}{2J+1}\right)$ versus E_J/k, the slope of the line will be $-1/T_{\rm ex}$ and the y-intercept will be $\ln\left(\frac{N_T}{Z(T_{\rm ex})}\right)$. If the assumption of constant $T_{\rm ex}$ is invalid, no straight line will fit the data. [show Manoj+ paper on Orion]

3.3.3 τ, N, and M

If we have observations of optically thin isotopologues, we can use their line ratios to derive the optical depth of the optically thin lines, and therefore the column density and cloud masses.

We are of course interested in the total column density of CO molecules. Use Boltzmann equation to get:

$$N_{\rm T} = N_J \frac{\sum_J (2J+1)e^{-hBJ(J+1)/kT_{\rm ex}}}{(2J+1)e^{-hBJ(J+1)/kT_{\rm ex}}}\,. \tag{3.52}$$

So if we can derive N_J, we will get our answer (assuming $T_{\rm ex}$ is known or can be assumed).

We know there is a relationship between the optical depth and n_J:

$$\tau_\nu = \int_S n_J \frac{g_J}{g_{J-1}} \frac{c^2 A_{ul}}{8\pi\nu^2}(1 - e^{-hBJ(J+1)/kT_{\rm ex}})\phi_\nu ds\,, \tag{3.53}$$

which integrated along the line of sight gives us

$$\tau_\nu = N_J \frac{g_J}{g_{J-1}} \frac{c^2 A_{ul}}{8\pi\nu^2} (1 - e^{-hBJ(J+1)/k\,T_{\rm ex}}) \phi_\nu \,, \tag{3.54}$$

where $N_J = \int n_J ds$ is the column density of molecules in the upper state and $\phi(\nu)$ is the normalized line profile, usually taken to be a Gaussian. Or integrated over the line,

$$N_J = \frac{g_{J-1}}{g_J} \frac{8\pi\nu^2}{c^2 A_{ul}} \frac{1}{1 - e^{-hBJ(J+1)/k\,T_{\rm ex}}} \int \tau_\nu d\nu \tag{3.55}$$

so finally

$$N_T = \frac{8\pi\nu^2}{A_{ul} c^2} \frac{\Sigma_J (2J+1) e^{-hBJ(J+1)/kT_{\rm ex}}}{(2J+1) e^{-hBJ(J+1)/kT_{\rm ex}}} \int_{-\infty}^{\infty} \tau_\nu d\nu \tag{3.56}$$

We observe ΔT_B from Equation 3.44. Solving for τ_ν, we find

$$\tau_\nu = -\ln \left[1 - \frac{\Delta T_B}{T_0} \left(\frac{1}{e^{T_0/\,T_{\rm ex}} - 1} - \frac{1}{e^{T_0/T_{\rm BG}} - 1} \right)^{-1} \right] , \tag{3.57}$$

where $T_0 = h\nu/k$. For ^{13}CO $J = 1 \rightarrow 0$, if we assume $T_{\rm BG} = 2.7$ K, this is

$$\tau_0^{\rm 13CO} = -\ln \left\{ 1 - \frac{\Delta T_B}{5.3} \left[\left(e^{5.3/T_{\rm ex}} - 1 \right)^{-1} - 0.16 \right]^{-1} \right\} . \tag{3.58}$$

Putting it all together,

$$N_T = \frac{8k\pi\nu^2}{A_{ul} h c^3} \frac{\Sigma_J (2J+1) e^{-hBJ(J+1)/kT_{\rm ex}}}{(2J+1) e^{-hBJ(J+1)/kT_{\rm ex}}} \int_{-\infty}^{\infty} -\ln \left[1 - \frac{\Delta T_B}{T_0} \left(\frac{1}{e^{T_0/\,T_{\rm ex}} - 1} - \frac{1}{e^{T_0/T_{\rm BG}} - 1} \right)^{-1} \right] d\nu \tag{3.59}$$

At this point we have everything we need, but should probably simplify Equation 3.59. It is important to point out though that Equation 3.59 is a perfectly acceptable solution, with one measured quantity (ΔT_B), one derived ($T_{\rm ex}$), and one assumed ($T_{\rm BG}$). To make it simpler,

Assumption #1: $Z(T_{\rm ex}) \simeq [1 + (k\,T_{\rm ex}/hB)^2]^{1/2}$, or even that $Z(T_{\rm ex}) \simeq (k\,T_{\rm ex}/hB)$. This makes the left hand side of Equation 3.59 reduce to something reasonable. For ^{13}CO again, using the latter expression,

$$N_{T,\rm 13CO} = 3 \times 10^{14} \frac{T_{\rm ex} \int_{-\infty}^{\infty} \tau_{13CO} d\nu}{1 - e^{-5.3/\,T_{\rm ex}}} \,. \tag{3.60}$$

Assumption #2: Working with τ_ν is not ideal, since our expression for it is so ugly. If we

have an optically thin tracer, $T\tau_\nu \simeq \Delta T_B$, but only approximately so. A better relationship is:

$$T \int_{-\infty}^{\infty} \tau_\nu d\nu \simeq \frac{\tau_\nu}{1 - e^{-\tau_\nu}} \int_{-\infty}^{\infty} \Delta T_B d\nu \,. \tag{3.61}$$

Rohlfs & Wilson state that this is accurate to within 15% for $\tau_0 < 2$ and always overestimates N_T for $\tau_0 > 1$. So using Assumption #1:

$$N_{T,13\mathrm{CO}} = 3 \times 10^{14} \frac{\frac{\tau_\nu}{1-e^{-\tau_\nu}} \int_{-\infty}^{\infty} \Delta T_B d\nu}{1 - e^{-5.3/\,T_{\mathrm{ex}}}} \,. \tag{3.62}$$

Assumption #3: If τ_ν is very low, $\frac{\tau_\nu}{1-e^{-\tau_\nu}} \simeq 1$ so

$$N_{T,13\mathrm{CO}} = 3 \times 10^{14} \frac{\int_{-\infty}^{\infty} \Delta T_B d\nu}{1 - e^{-5.3/\,T_{\mathrm{ex}}}} \,. \tag{3.63}$$

If we can assume $T_{\mathrm{ex}} = 10\,\mathrm{K}$ and Gaussian line profiles,

$$N_{T,13\mathrm{CO}} = 8.75 \times 10^{14} T_B \Delta v. \tag{3.64}$$

Sometimes ^{13}CO is optically thick, and we need to substitute in C^{18}O, or use the $J = 2 \rightarrow 1$ transition.

This is known as the LTE treatment where all isotopologues have the same value of T_{ex} and the Boltzmann equation applies. LTE is commonly assumed, and is generally required to make the problem tractable. It is an assumption though!

The problem is that we don't really care about CO because it exists in relatively low quantities; we want to know about H_2. To convert between these two, we have the X_{CO} factor. X_{CO} is known to be around 2×10^{20}, but varies by environment and by metallicity (see excellent review by Bolatto, also Draine 19.6). Because of uncertainties in X_{CO}, the values derived from the use of CO to trace the total molecular column is also uncertain.

[Cloud Masses (Simon et al. 2001)] How can we use the above equation? What we are generally interested in is cloud masses. *Assuming we have observations of some optically thin tracer, we have column densities, and therefore can easily integrate over the cloud to get a mass.* We have to assume a ^{12}CO to ^{13}CO ratio, and an X_{CO} value, in addition to assuming LTE. We are finally left with

$$M_{\mathrm{LTE}} = 0.96\, M_\odot \left[\frac{N_{T,13\mathrm{CO}}}{8.75 \times 10^{14} \mathrm{cm}^{-2}}\right] \left(\frac{\theta_x}{60''}\right) \left(\frac{\theta_y}{60''}\right) \left(\frac{D}{\mathrm{kpc}}\right)^2 \tag{3.65}$$

this from Simon et al. (2001), where cloud dimensions are θ and cloud distance D.

It can be much more complicated! Check out Mangum & Shirley (2017).

3.3.4 Results of CO observations

[Observations of molecular clouds (Draine Ch. 32)]
Molecular gas is $\sim 22\%$ of all mass in the Milky Way, $M(\mathrm{H_2}) \simeq 8.4 \times 10^8\ M_\odot$. Molecular clouds form stars. We separate clouds into categories based on their optical appearance in terms of optical depth (Draine Table 32.1). These are "diffuse molecular clouds" with $A_V \lesssim 1$, "translucent clouds" with $A_V \simeq 1-5$, "dark clouds" with $A_V \simeq 5-20$ and "infrared dark clouds" with $A_V \simeq 20 -> 100$. Remember your definitions for A_V!

Extinction is wavelength dependent such that longer wavelengths pass more readily through material. Therefore, at short wavelengths, all clouds are observed in absorption. As we go toward longer wavelengths, we see clouds in emission.

We may also classify clouds based on their sizes and masses (Draine Table 32.2). These range from "GMC complexes" (GMC=giant molecular cloud) to cores. Cores have the smallest sizes, the highest densities, the smallest line widths (and therefore the smallest temperatures), the smallest masses, and the highest A_V.

A GMC complex is a gravitationally bound complex of GMCs. Orion is the nearest. GMCs themselves have lots of structure. Clumps can actually be isolated - we need not have all this structure.

The number of GMCs in the Galaxy can be estimated with

$$\frac{dN_{\rm GMC}}{d\ln M_{\rm GMC}} = N_u \left(\frac{M_{\rm GMC}}{M_u}\right)^{-\alpha} \tag{3.66}$$

for $10^3\ M_\odot \lesssim M_{\rm GMC} < M_u$, with $M_u \simeq 6 \sim 10^6\ M_\odot$, $N_u \simeq 63$, and $\alpha \simeq 0.6$ (Williams & McKee, 1997). Notice the similarity to the IMF! Nature loves power laws. For this distribution, most of the mass is in the most massive GMCs, $> 80\%$ of the total mass in GMCs with masses $> 10^5\ M_\odot$.

3.3.5 Summary of CO Use

Q: How can we derive $T_{\rm ex}$?
A: 1) Observe an optically thick tracer. In this case, $\Delta T_B \simeq T_{\rm ex} \simeq T_k$
A: 2) Observe multiple transitions of the same tracer, and fit the rotational energy diagram (Equation 3.51).

Q: How can we derive the column density (and mass)?
A: *We need an optically thin tracer!* Then, we need to determine the excitation temperature. Then, the line integrated intensity (in K $\mathrm{km\,s^{-1}}$) is proportional to the column density with some assumptions. We get the mass from the cloud shape.

3.4 Symmetric Tops, or Symmetric Rotors, NH_3 (Draine 5.2.1)

The next-most-complicated molecules are called "symmetric tops" or "symmetric rotors."

In general, molecular energies are given by:

$$E_r = AhJ(J+1) + BhJ(J+1) + ChJ(J+1)\,, \tag{3.67}$$

where the rotational constants A, B, and C are defined as before, e.g., $B = \frac{h}{8\pi^2 I_B}$. The change here from our treatment of CO is that now I_B is the projection of the angular momentum onto the B axis.

Ammonia is a *symmetric top* molecule, which means that it is symmetric under rotation ($I_B = I_C$). The energy of symmetric top transitions is

$$E = hBJ(J+1) + h(A-B)K^2\,, \tag{3.68}$$

where K is the projection of J along the symmetry axis (K is integer from 0 to J), and A and B are rotational constants.

Ammonia also has *inversion* transitions. These are different from rotational transitions. When $J = K$, the levels are metastable (most molecules found with $J = K$) and the molecule is spinning around its symmetry axis. These are known as the "rotational backbone." Classically, there is no electric dipole radiation possible.

Inversion transitions arise from the two distinct eigenstates for the NH_3 wavefunction. This can be thought of as the N atom tunneling through the triangle formed by the three H atoms. Inversion transitions take the compact notation (J, K), e.g., (1,1), although the transition is between two different energy states, each with the same (J, K). Transitions are near 23 GHz.

The inversion transitions are further split by interactions with the electric quadrupole moment of the nitrogen nucleus and the magnetic dipole of the protons. The (1,1) line for example is split into four separate hyperfine lines.

This splitting into hyperfine states is confusing, but it occurs all the time. Even familiar spectral lines are frequenctly composed of multiple distinct hyperfine transitions. Often, these transitions are close to each other with respect to the line width, and so form a single spectral line. For example, take a look at figures 6 and 7 in the Magnum paper.

Just like for CO, we can use NH_3 to derive mass and optical depth. Line ratios of the various inversion lines give you the optical depth. I invite you to look at the excellent review of NH_3 by Ho & Townes.

[Kwok 7.9] Formaldehyde (H_2CO) is a near-symmetric rotor. The 6 cm $1_{11} \rightarrow 1_{10}$ transition was one of the first transitions detected in the radio (in 1969). This line is (almost) always seen in absorption. This suggests that formaldehyde is an "anti-maser," where the upper level is depopulated (really low $T_{\rm ex}$).

3.5 Asymmetric tops, or Asymmetric Rotors, H_2O (Draine 5.2.2)

Asymmetric tops are much more complicated.... H_2O does have a rotational backbone, just like NH_3. Just like for NH_3, the backbone states are highly populated. For most backbone states, the only allowed transitions are to other backbone states. States are specified as $J_{K-1,K+1}$. H_2O has two independent ladders, with total nuclear spin 0 ("para") or 1 ("ortho"). There are no transitions between para and ortho. The selection rules for electric dipole radiation are $\Delta J = 0, \pm 1; \Delta K_{-1} = \pm 1, \pm 3; \Delta K_{+1} = \pm 1, \pm 3$.

There are two ortho and two para transitions for which the backbone states can have transitions to non-background levels (ortho: $4_{14} \rightarrow 3_{21}$ and $6_{16} \rightarrow 5_{23}$; para $3_{13} \rightarrow 2_{20}$ and $5_{15} \rightarrow 4_{22}$). We can see these as masers. The upper level is over populated compared to the lower. This is a population inversion.

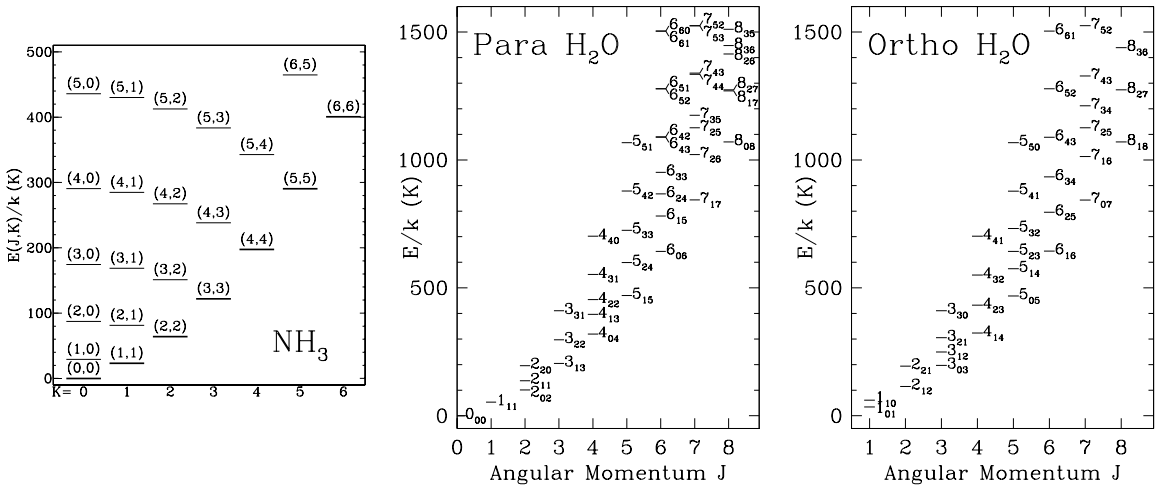


Figure 3.4: Energy level diagrams for NH_3 and H_2O. The rotational backbone for NH_3 are those states with $K_a = K_c$. H_2O masers can be found for transitions ortho: $4_{14} \rightarrow 3_{21}$ and $6_{16} \rightarrow 5_{23}$; para $3_{13} \rightarrow 2_{20}$ and $5_{15} \rightarrow 4_{22}$.

3.6 OH and "Λ Doubling"

Hydroxyl (OH) is a bit of an oddball. Because it is such a simple molecule, it is abundant in the ISM. OH has 7 electrons, so the lowest electron orbital momentum is $\mathcal{L} = 1$ and spin $\vec{\Sigma} = 1/2 * 2 + 1 = 2$, with $\vec{J} = 1 - 1/2, 1 + 1/2$. This state is therefore

$$^2\Pi_{1/2,3/2}\,. \tag{3.69}$$

This gives rise to independent rotational ladders for each state ($J = 1/2, J = 3/2$) due to spin-orbit coupling.

OH is super interesting because each of these rotational states are split by two different mechanisms. First, each is split by "Λ-doubling," an interaction between the nuclear rotation

and the unpaired electron motion around its orbit. There is a slight splitting of the energy levels because the rotational momentum of the molecule is very slightly transferred to the orbital electrons, causing them, in effect, to become excited to a higher state. Given two molecules, one with an electron cloud aligned with the rotation axis, and one that will not, the molecule with the aligned cloud will have a smaller moment of inertia and therefore lower energies.

Secondly, there are "hyperfine" splittings of each Λ-doubling split level. OH has non-zero nuclear spin, and therefore has a nuclear magnetic moment. The magnetic field from the nuclear magnetic moment couples to the electron motions and therefore the electronic energy depends on the orientation of the nuclear moment and the electron angular momentum. The hyperfine lines are from quantum number $F = J + S$.

For OH, $S = 1/2$, and for $J = 3/2$ we have $F = 1, 2$. Thus, there are four energy levels associated with the $^2\Pi_{1/2,3/2}$ state: two from Λ-doubling which are then split by the hyperfine interaction. Transitions between these four states are at 1720.5, 1667.4, 1665.4, and 1612.2 MHz.

OH is now studied extensively, because researchers think that it is key to understanding so-called "dark gas." We know that CO is missing some molecular material, because it has a relatively high critical density. For OH, the critical density is much lower, so it is an excellent probe of the diffuse ISM.

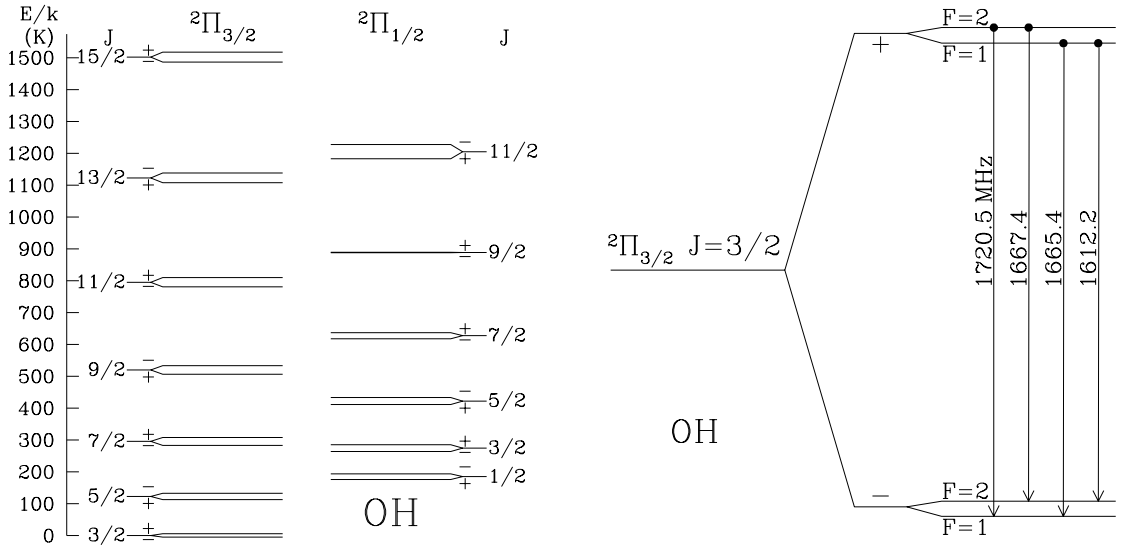


Figure 3.5: OH energy level diagram showing Λ-doubling and hyperfine splitting.

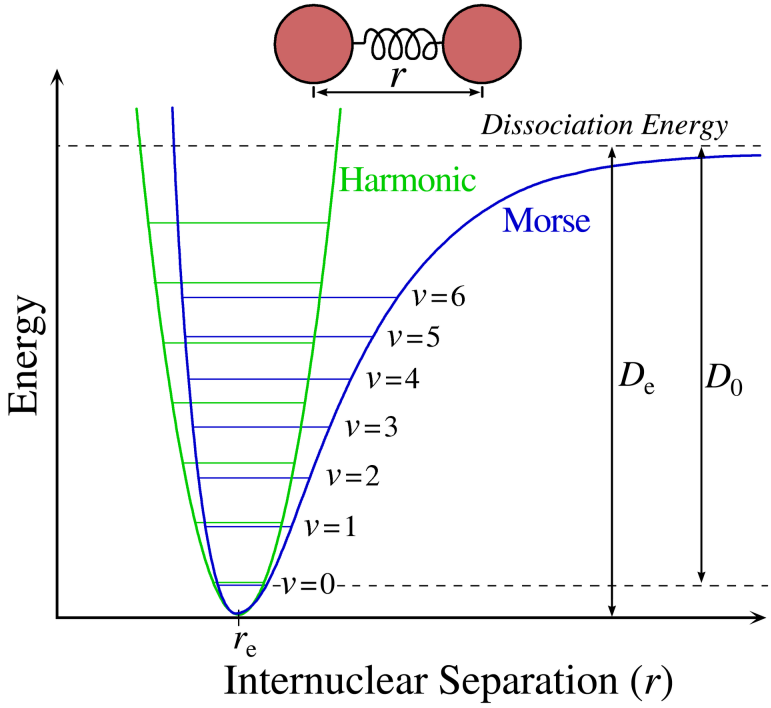


Figure 3.6: The Morse potential (from Wikipedia).

3.7 Vibration

At the beginning of our discussion on molecules, we learned that molecules have three important types of energies that give rise to radiation: electronic, vibrational, and rotational. We derived characteristic energies for these transitions and found that electronic transitions are in the optical/UV regime, vibrational in the IR, and rotational in the radio.

We can think of vibrational, or stretching, energies as being due to two atoms attached with a spring. What does the potential look like in this case? It is basically just the classical solution: that of a simple harmonic oscillator (SHO), with energies $E \propto r^2$, where r is again the difference in distance between the atoms. The actual potential is can be modeled as the so-called "Morse Potential":

$$E(R) = E_0 \left[1 - e^{-(r-r_0)/L}\right]^2 , \tag{3.70}$$

where L is the angular momentum. Notice that this is a central potential that only depends on the distance between the two atoms r. This potential is at a minimum when $r = r_0$, the equilibrium distance. We can expand the Morse potential to get:

$$E(r) = E_0(r_0) + 1/2\mu\omega_0^2(r - r_0)^2 + \dots , \tag{3.71}$$

where $\omega_0 = (k/\mu)^{1/2}$ for the SHO, where k is the spring constant (2nd term $= 1/2k(r-r_0)^2$). The spring constant k is closely related to the strength of the chemical bond. These terms are actually rather simple! We have a constant plus a SHO potential. The physical meaning of these terms are just:

$$E(r) = E_{\rm bind} + E_{\rm vibrational} , \tag{3.72}$$

the binding and vibrational energies of the molecule.

We are interested here in the vibrational energies. For the SHO, the solutions are Hermite polynomials: [from Kwok section 7.5]

$$E_v = (v + 1/2)h\nu_0 \,, \tag{3.73}$$

where v is the vibrational quantum number and

$$\nu_0 = \frac{1}{2\pi}\sqrt{\frac{k}{\mu}} \tag{3.74}$$

is the natural oscillator frequency, k is the spring constant. We immediately arrive at our first result: the vibrational frequency decreases with increasing atomic mass. Heavier molecules will have lower vibrational frequencies. The approximations above break down for large values of v. In this case, the Morse potential stops looking parabolic. At high v, the energies get closer together than the SHO (smaller than energies given above), because the potential is flatter. Molecular hydrogen has 14 vibrational states before the continuum.

The selection rule for vibrational transitions within the harmonic approximation is $\Delta v = \pm 1$. The ground state transition $v = 1 \leftrightarrow 0$ is the "fundamental transition." Transitions from higher excited states to the ground state ($v = n \leftrightarrow 0, n = 2, 3, 4, ...$) are referred to as *overtones*. Transitions between excited states are called *hot bands*.

3.8 All together now!

We have now independently derived rotational and vibrational energies. These two couple to give the total energy for molecules:

$$E_{v,J} = \left(v + \frac{1}{2}\right) h\nu_0 - x\nu_0 + hBJ(J+1) - hD[J(J+1)]^2 + hH[J(J+1)]^3 + ... \,, \tag{3.75}$$

First term: Vibrational energy of SHO
Second term: Anharmonic term of deviation from pure SHO
Third term: Rotational energy
Fourth term: Centrifugal distortion term to account for the fact that as the rotation rate increases, the atoms get further apart, increasing the moment of inertia.
Fifth term: Rotational-vibrational coupling. Vibrations change the moment of inertia. Large vibrations increase the moment of inertia, and therefore decrease the energy.

There are constants associated with each of these, and these "constants" depend on the vibrational quantum number v.

3.8.1 Actual Emission

Rotational transitions within each vibrational transition are organized into branches according to the change in the rotational quantum number. The selection rule for one-photon

electric-dipole transitions is $\Delta J = -1, +1$. The $\Delta J = -1$ branch is known as the P-branch and the $\Delta J = +1$ branch is known as the R-branch. For the fundamental mode therefore

$$\nu = \nu_0 + 2(J+1)B; J = 0, 1, 2, ... \ \ (\mathrm{R-branch}) \tag{3.76}$$

and

$$\nu = \nu_0 - 2JB; J = 1, 2, 3, ... \ \ (\mathrm{P-branch})\,. \tag{3.77}$$

For electric quadrupole, the branches are O ($\Delta J = -2$) or S ($\Delta J = +2$). The Q-branch is $\Delta J = 0$, but this is not allowed when $J = 0$. The nomenclature of vibrational transitions is: "v branch (J_ℓ)". Therefore, the 1-0 S(1) transition of H_2 refers to the transition $(v = 1, J = 3) \rightarrow (v = 0, J = 1)$.

Draine Table 5.2

Designation	$(J_\mathrm{u} - J_\ell)$	Note
$O(J_\ell)$	-2	Electric quadrupole
$P(J_\ell)$	-1	Electric dipole
$Q(J_\ell)$	0	Electric dipole or electric quadrupole
$R(J_\ell)$	$+1$	Electric dipole
$S(J_\ell)$	$+2$	Electric quadrupole

It is a bit confusing how all this fits together. The energy level diagram in Figure 7.5 of Kwok or Figure 5.2 of Draine helps a lot. We can have transitions between rotational levels only, between vibrational levels only, or between both at the same time. Differences in vibrational quantum number v have the largest energies.

3.9 H_2 (finally!)

It is often said that H_2 cannot be directly observed. This is not true, it's just difficult to observe. Remember that driving cat on Saturday Night Live that would always get into accidents. It turns out he could drive, just not very well. H_2 is exactly like that in every way possible.

[from Pogge] There is no electric dipole for H_2 so we can't get electric dipole radiation. We can get electric quadrupole though. This means that only the $\Delta J = 0$ and $\Delta J = 2$ quadrupole transitions occur, while the $\Delta J = 1$ (dipole) rotational transitions are strictly forbidden. Unlike CO, H_2 emits no long-wavelength pure rotational lines. The P and R branches of H_2 do not occur, but transitions in the O, Q, and S branches can occur and are observed in the near- and mid-infrared.

The small mass and small size of the H_2 molecule gives it a low moment of inertia. The first vibrational-rotational transition in the H_2 ground state is the v=0-0S(0) transition, which is the J=2 $\rightarrow$ 0 transition at 28.2 μm, a part of the spectrum unobservable from the ground due to water-vapor absorption in the atmosphere. Further, $h\nu/k = 514\,\mathrm{K}$ for this transition, very large relative to typical temperatures in giant molecular clouds (10-20 K). [This is the

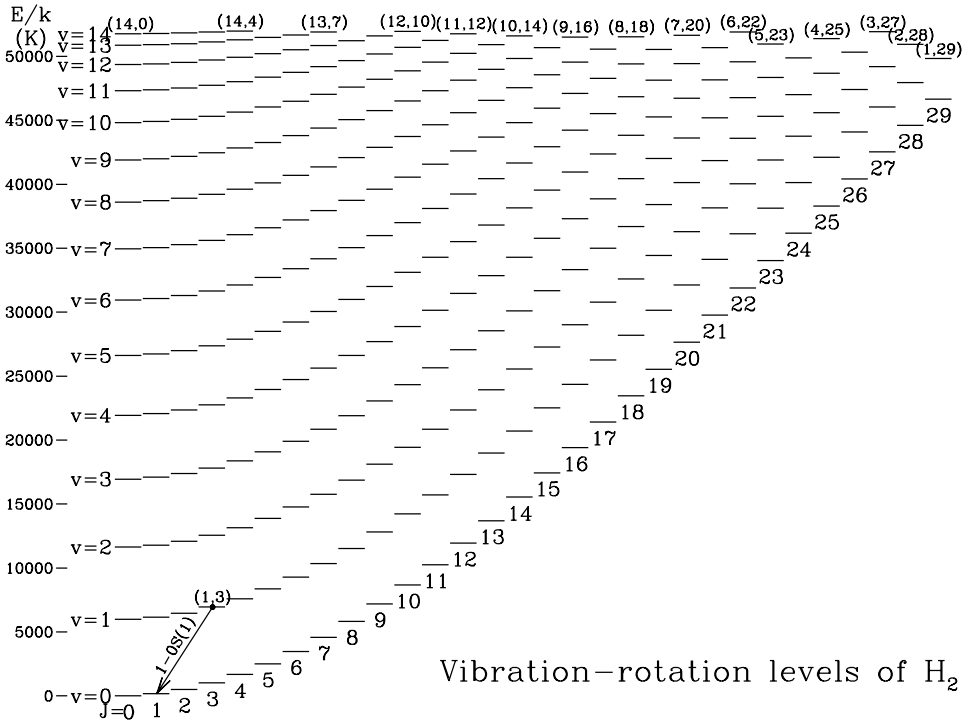
energy above ground state] All other vibrational-rotational transitions in the ground state have increasing energy (and shorter wavelength). The next pure rotational transition is $J = 4 \rightarrow 2$ at $12\,\mu$m, which has $h\nu/k \simeq 1200\,$K. The generally high energies of the first excited states of H_2 means that we expect negligible H_2 emission unless we are looking at unusually warm (500-1000 K) H_2 gas in proximity to hot stars or in regions of active star formation within or at the fringes of giant molecular clouds. [H_2 dissociates at $\sim 50000\,$K] For a population in equilibrium at a temperature of 10 K, the fraction of molecules in the $J = 2$ state is $\sim e^{-10/500} \sim 10^{-22}$ In effect, in a molecular cloud there are simply no H_2 molecules in states capable of emitting.

Pure rotational lines ($v_{\rm u} = v_\ell$) span the 3.4 to $28\,\mu$m region, while $\Delta v = \pm 1$ vibrational-rotational transitions have typical energies of ~ 0.5eV and are found in the $1.4\,\mu$m region clustering around $2\,\mu$m. Typical transition probabilities are $\sim 10^{-7}$ to $10^{-8}\,{\rm s}^{-1}$, so these lines are strongly forbidden.

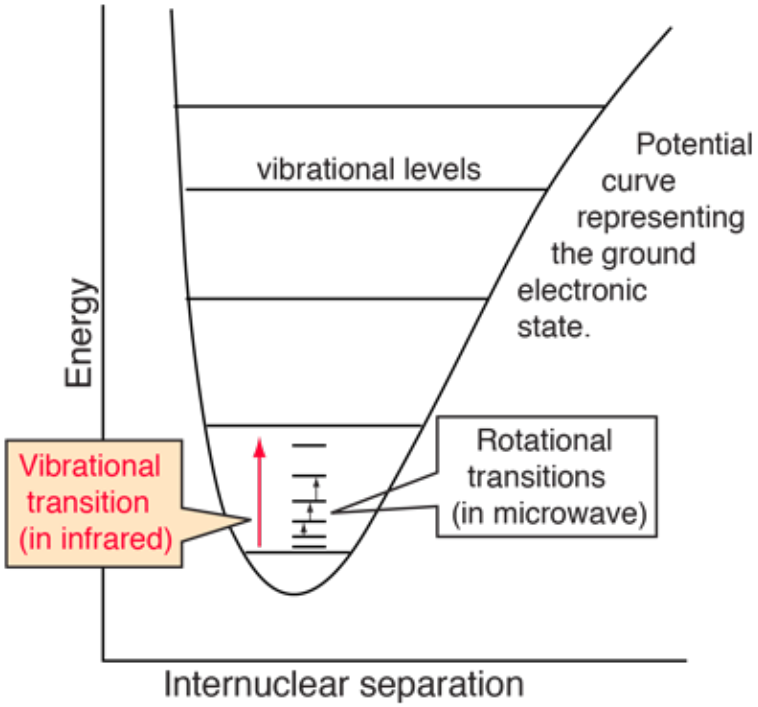
Contrast this situation with the CO $J = 1 \rightarrow 0$ line at 2.6 mm where $h\nu/k = 5.53\,$K and the molecule, which is easily excited by H_2 or H I collisions at temperatures of $T = 10 - 20\,$K more typical of the cores of giant molecular clouds. In general, H_2 is only directly observable as:
1. Absorption at Far-UV wavelengths in the diffuse ISM along sight lines toward nearby stars in the Lyman and Werner band electronic transitions. These lines arise in both cold and warm H_2. This is our only direct probe of the cold H_2 gas that makes up most of the ISM.
2. Emission by Infrared rotational-vibrational transitions in the electronic ground state of H_2 at wavelengths between 1 and 28μm in relatively warm regions. The molecular gas must be warm (500-2000 K), excited either by shocks, outflows, or UV fluorescence from nearby stars.

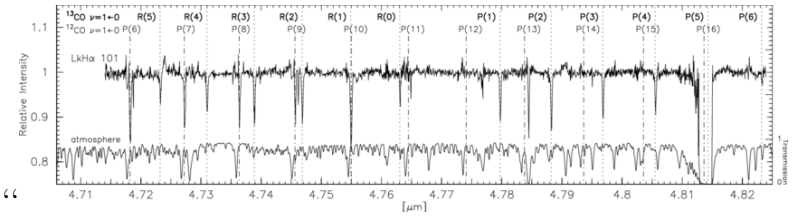
[Draine 5.1.6] One final note about H_2: there are ortho and para branches as we saw for H_2O. Again, this arises from the fact that the combined proton spins can be either anti-aligned (0, ortho) or aligned (1, para). Transitions cannot change the ortho/para branch. Like H_2O therefore the two branches operate like different species.

[Kwok] One super final note: HD is not homonuclear, and does have long wavelength rotational lines in the far infrared. These lines can be important.



Vibration–rotation levels of H_2




^{13}CO ν=1←0 R(5) R(4) R(3) R(2) R(1) R(0) P(1) P(2) P(3) P(4) P(5) P(6)
^{12}CO ν=1←0 P(6) P(7) P(8) P(9) P(10) P(11) P(12) P(13) P(14) P(15) P(16)
LkHα 101
atmosphere
Relative Intensity
1.1
1
0.9
0.8
4.71 4.72 4.73 4.74 4.75 4.76 4.77 4.78 4.79 4.8 4.81 4.82
[μm]

“

Chapter 4

Ionized Gas (and Some Atomic)

First, a review:
In the introduction to this course, we talked about different phases of the ISM. Together, ionized gas makes up about 23% of the mass of the ISM in our Galaxy. This total is broken down into Coronal gas that has $T \gtrsim 10^{5.5}\,\mathrm{K}$, $n \simeq 0.004\ \mathrm{cm}^{-3}$ (the "hot ionized medium") and gas with $T \gtrsim 10^{4}\,\mathrm{K}$, $n \simeq 0.3 - -10^{4}\ \mathrm{cm}^{-3}$ (discrete H II regions and the "Warm Ionized Medium"). Of these, the WIM contains the most mass.

Obviously our diagnostics of the ionized gas will be temperature sensitive: certain transitions will be detected in the coronal gas, and a different set of transitions will be detected in the WIM. Keep this in mind as we review quantum mechanics and selection rules.

4.1 Ionization Processes

Some gas is ionized, and some is not. How about that?
In dense molecular clouds, the material is essentially all neutral, with $n_e/n_H \lesssim 10^{-6}$. Ionizing radiation cannot penetrate. Ionizations due to cosmic rays (CRs).
In H I C is frequently ionized (low ionization potential, remember), and H I is somewhat ionized by CRs., $10^{-3} \lesssim n_e/n_H \lesssim 10^{-1}$, depending on density, temperature, CR ionization rate.
In an H II region, hydrogen completely ionized. He has an ionization potential of $\sim 24\,\mathrm{eV}$, and so may be ionized. C, O, N ionized, and OIII and NIII may be present as well, depending on hardness of radiation field.
In the HIM, carbon is in CIV or CV, and oxygen in OIV, OV, or maybe OVI.

The two main ionization processes are photoelectric absorption, and collisional ionization. Cosmic ray ionization can also be important, as can charge exchange.

Appendix D

Ionization Potentials (eV)

	Element	I→II	II→III	III→IV	IV→V	V→VI	VI→VII	VII→VIII
1	H	13.5984						
2	He	24.5874	54.416					
3	Li	5.3917	75.640	122.454				
4	Be	9.3227	18.211	153.894	217.719			
5	B	8.2980	25.155	37.931	259.375	340.226		
6	C	11.2603	24.383	47.888	64.494	392.089	489.993	
7	N	14.5341	29.601	47.449	77.474	97.890	552.072	667.046
8	O	13.6181	35.121	54.936	77.414	113.899	138.120	739.293
9	F	17.4228	34.971	62.708	87.140	114.243	147.163	185.189
10	Ne	21.5645	40.963	63.423	97.117	126.247	154.214	207.271
11	Na	5.1391	47.286	71.620	98.91	138.40	172.183	208.50
12	Mg	7.6462	15.035	80.144	109.265	141.270	186.76	225.02
13	Al	5.9858	18.829	28.448	119.992	153.825	190.477	241.76
14	Si	8.1517	16.346	33.493	45.142	166.767	205.267	246.481
15	P	10.4867	19.769	30.203	51.444	65.025	220.422	263.57
16	S	10.3600	23.338	34.790	47.222	72.594	88.053	280.948
17	Cl	12.9676	23.814	39.911	53.465	67.819	97.030	114.201
18	Ar	15.7596	27.630	40.735	59.686	75.134	91.00	124.328
19	K	4.3407	31.628	45.806	60.913	82.66	99.4	117.6
20	Ca	6.1132	11.872	50.913	67.27	84.51	108.8	127.2
21	Sc	6.5615	12.800	24.757	73.489	91.69	110.7	138.0
22	Ti	6.8281	13.576	24.492	43.267	123.7	119.533	140.846
23	V	6.7462	14.655	29.311	46.709	65.282	128.125	150.641
24	Cr	6.7665	16.486	30.959	49.160	69.456	90.635	160.175
25	Mn	7.4340	15.640	33.668	51.2	72.4	95.60	119.203
26	Fe	7.9024	16.188	30.651	54.801	75.010	99.063	124.976
27	Co	7.8810	17.084	33.50	51.27	79.5	102.	129.
28	Ni	7.6398	18.169	35.187	54.925	76.06	107.87	133.
29	Cu	7.7264	20.292	36.841	57.380	79.846	103.031	138.862
30	Zn	9.3492	17.964	39.723	59.573	82.574	133.903	133.903

Notes:

- Ionization potentials from Ralchenko et al. (2010).
- The light line separates ions with $I < I_{\rm IIe}$ from ions with $I > I_{\rm IIe} = 24.6\,{\rm eV}$.
- Ions to right of the heavy line (with $I > I_{\rm IIe\,II} = 54.4\,{\rm eV}$) are not abundant in gas photoionized by O or B stars and are therefore indicative of photoionization by WR stars, PN nuclei, or collisional ionization in shocked gas.
- For elemental abundances, see Table 1.4.

Figure 4.1: Ionization potentials from Draine Appendix D

4.1.1 Collisional Ionization

The rate of collisional ionization in $\mathrm{cm}^{-3}\ \mathrm{s}^{-1}$ is

$$k_{\mathrm{ci}} = \int_I^\infty \sigma_{\mathrm{ci}}(E) v f_E dE \,, \tag{4.1}$$

This makes perfect sense, actually. We know that collision rate is $C = n\sigma v$, so $C/n = \sigma v$. The rms velocity of a maxwellian distribution is $\langle v \rangle = \int_0^\infty v f(v)\, dv$, but we are only interested in the range above the ionization energy I. For Maxwellian distributions,

$$k_{\mathrm{ci}} = \left(\frac{8kT}{\pi m_e}\right)^{1/2} \int_I^\infty \sigma_{\mathrm{ci}}(E) \frac{E}{kT} e^{-E/kT} \frac{dE}{kT} \,. \tag{4.2}$$

We can approximate the collisional cross section:

$$\sigma_{\mathrm{ci}} \approx C\pi a_0^2 \left(1 - \frac{I}{E}\right) \,, \tag{4.3}$$

where C is a constant of order unity. This expression reduces to zero when $I = E$. This gives us our expression for the rate coefficient:

$$k_{\mathrm{ci}} = 5.466 \times 10^{-9}\, C T_4^{1/2} E^{-I/kT}\ \mathrm{cm}^{-3}\,, \mathrm{s}^{-1} \,. \tag{4.4}$$

We see that as the temperature increases, the collisional ionization rate increases slowly. If the ionization potential is higher, the rate drops off rapidly.

4.1.2 Photoionization

The photoionization rate in s^{-1} is

$$\zeta_{\mathrm{p.i.}} = \int_{\nu_1}^\infty \sigma_{\mathrm{pe}}(\nu) c \frac{u_\nu}{h\nu}\, d\nu \,, \tag{4.5}$$

where ν_1 is the frequency of radiation required to ionize the element. Draine gives expressions for σ_{pe}. Notice that at high energies, the "metals" dominate the photoionization cross sections. Even though the abundance of metals is low, their impact is really large.

4.1.3 Cosmic Ray Ionization

Cosmic rays are energetic electrons and ions. Their velocities are much greater than those drawn from a thermal distribution, and thus they are "non-thermal." The mean kinetic energy of CRs is $\approx 35\,\mathrm{eV}$ according to Draine.

Without too much work, we can come up with an expression for the cosmic ray ionization rate:

$$\zeta_{\mathrm{CR}} = 4\pi \int_{E_{\mathrm{min}}}^\infty \sigma_{\mathrm{ci}}(E) \frac{dF}{dE} \frac{dE}{E} \,, \tag{4.6}$$

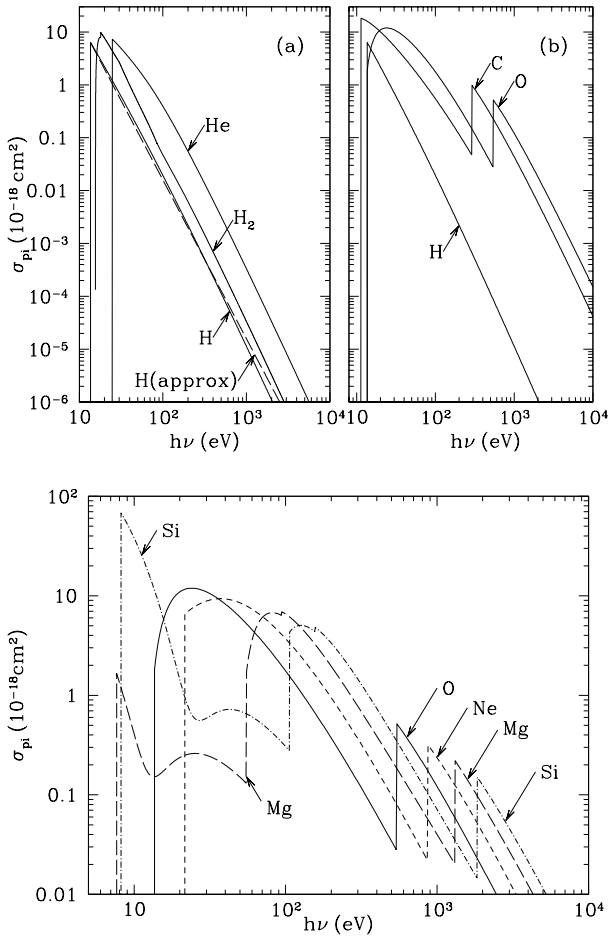


Figure 4.2: Photoionization cross sections for H, H_2, He, C, and O. Jumps for C and O correspond to ionizations from the K-shell ($1s$).

where $\sigma_{\rm ci}$ is the cosmic ray ionization cross section, the flux per solid angle is $F(E)$, and $E_{\rm min}$ is the minimum energy required.

Draine gives an expression for $\sigma_{\rm ci}$. I don't find the physics terribly enlightening. CR ionization is important in high density environments, however, like dense dark clouds. Here, CR ionization provides the only mean of generating ions.

Draine lists in Chapter 16 that $\zeta_{\rm CR} = 1 \times 10^{16}\,{\rm s}^{-1}$.

4.2 Ionization in Predominantly Neutral Regions [Draine Chapter 16]

The subject of CR ionization leads directly into our next topic, of ionization in neutral regions. Draine lists three distinct regimes:
Diffuse H I regions, where metals are photo-ionized by starlight. Remember, the ionization

cross section for metals greatly exceeds that for H. The gas may be $\approx 10^2$ K (CNM) or warm, ≈ 5000 K (WNM).
Diffuse molecular clouds (A_V between ~ 0.3 and 2). Most of the hydrogen is molecular, but metals still photoionized by starlight. CRs produce H_2^+, which leads to the formation of H_3^+.
Dark molecular clouds ($A_V \gtrsim 3$). There is insufficient photon flux to ionized C and S. CRs maintain a small fractional ionization $n_e/n_H \approx 10^{-7}$.

4.2.1 Ionization of Metals in H I Regions

Carbon is a very interesting element because it is abundant ($n_C/n_H \approx 10^{-4}$ after taking into account depletion in solid grains), and has an ionization potential less than H.

How much of the carbon is ionized?

If we have an ionization rate of carbon $\zeta(C^0)$ due to primarily photoioinization by starlight (with contributions from X-rays and CRs). The percentage ionized will be determined between the balance between ionization and recombination. The recombination rate must depend on the number of free electrons available, which from studies of pulsar DMs is $n_e \approx 0.04$ cm^{-3}. For a neutral fraction $x(C^0) = n(C^0)/n_C$, and recombination rate in cm^{-3} s^{-1} of $\alpha_{\rm rr}$,

$$\alpha_{\rm rr}(C^+) n_e [1 - x(C^0)] = \zeta(C^0) x(C^0) \tag{4.7}$$

which leads to

$$x(C^0) = \frac{\alpha_{\rm rr}(C^+) n_e}{\alpha_{\rm rr}(C^+) n_e + \zeta(C^0)} \, . \tag{4.8}$$

We can estimate most of these values: $\zeta(C^+) = 2.58 \times 10^{-10}$ s^{-1} (Draine Table 13.1), $\alpha_{\rm rr} = 8.63 \times 10^{-12}$ cm^{-3} s^{-1} for $T = 100$ K (Draine Table 14.6). Thus $x(C^0) = 1.3 \times 10^{-3}$.

This is an amazing result! About 99.9% of gas-phase interstellar carbon is ionized.

We get similar results for other metals with ionization potentials less than 13.6 eV.

4.2.2 Ionization of H in H I Regions

In the CNM, starlight with photon energies between 13.6 and 54.4 eV cannot penetrate into the cloud (Draine Figure 13.1; Figure 4.2. These are the ionization potentials for H and He (2nd potential). Higher energy photons can penetrate, however (Draine Figure 13.1; Figure 4.2). The column density of the CNM is $\sim 10^{20}$, and a 150 eV photon has absorption cross section per H of $18./ \times 10^{-20}$ cm^{-3} (due mostly to He). X-rays can have a big effect on the ionization of the CNM, if there is an x-ray source. CR ionization is dominant where x-rays are not.

Draine works out the estimate for the ionized fraction $x_e = n_e/n_{\rm H}$, in Figure 16.1.

The result is that the ionization fraction of hydrogen in the CNM is low, between 10^{-3} and 10^{-4}. The result is dependent on the CR ionization rate, which is pretty uncertain.

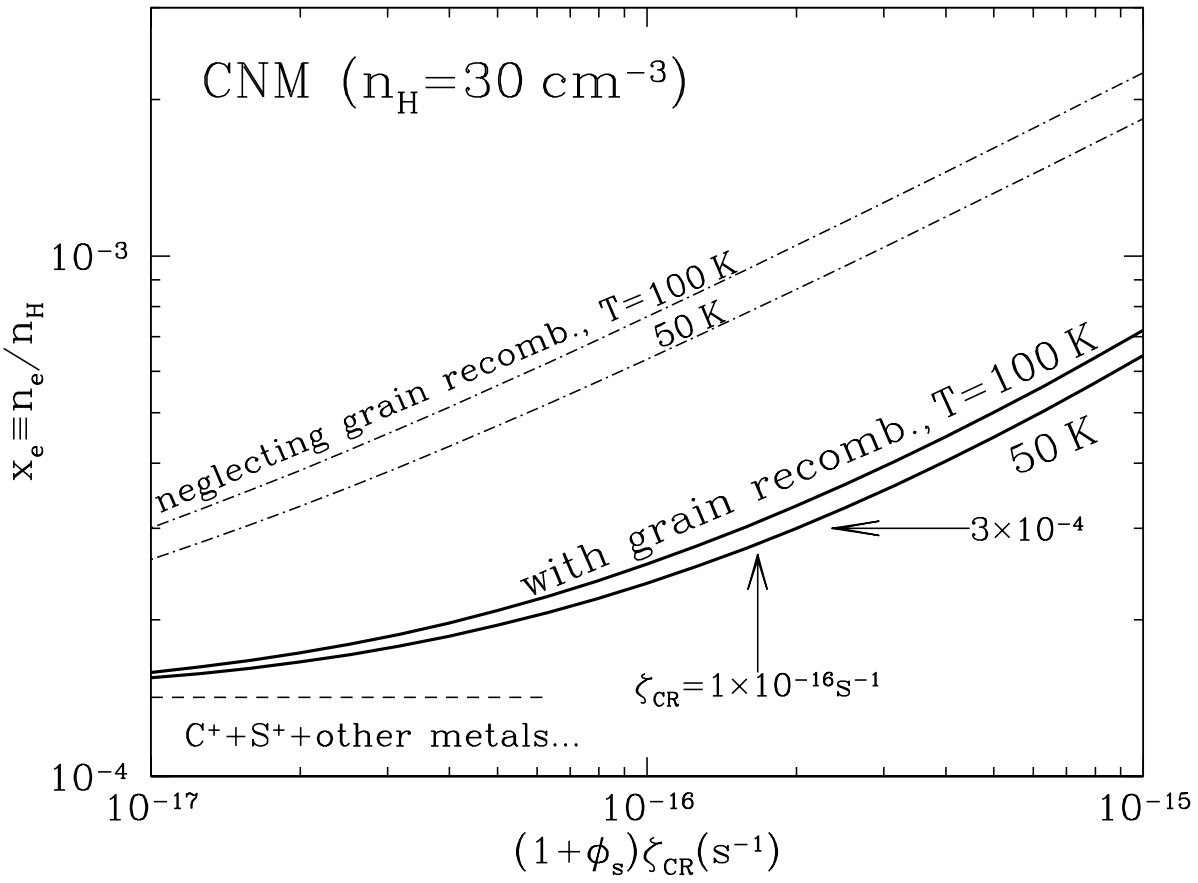


Figure 4.3: Draine 16.1

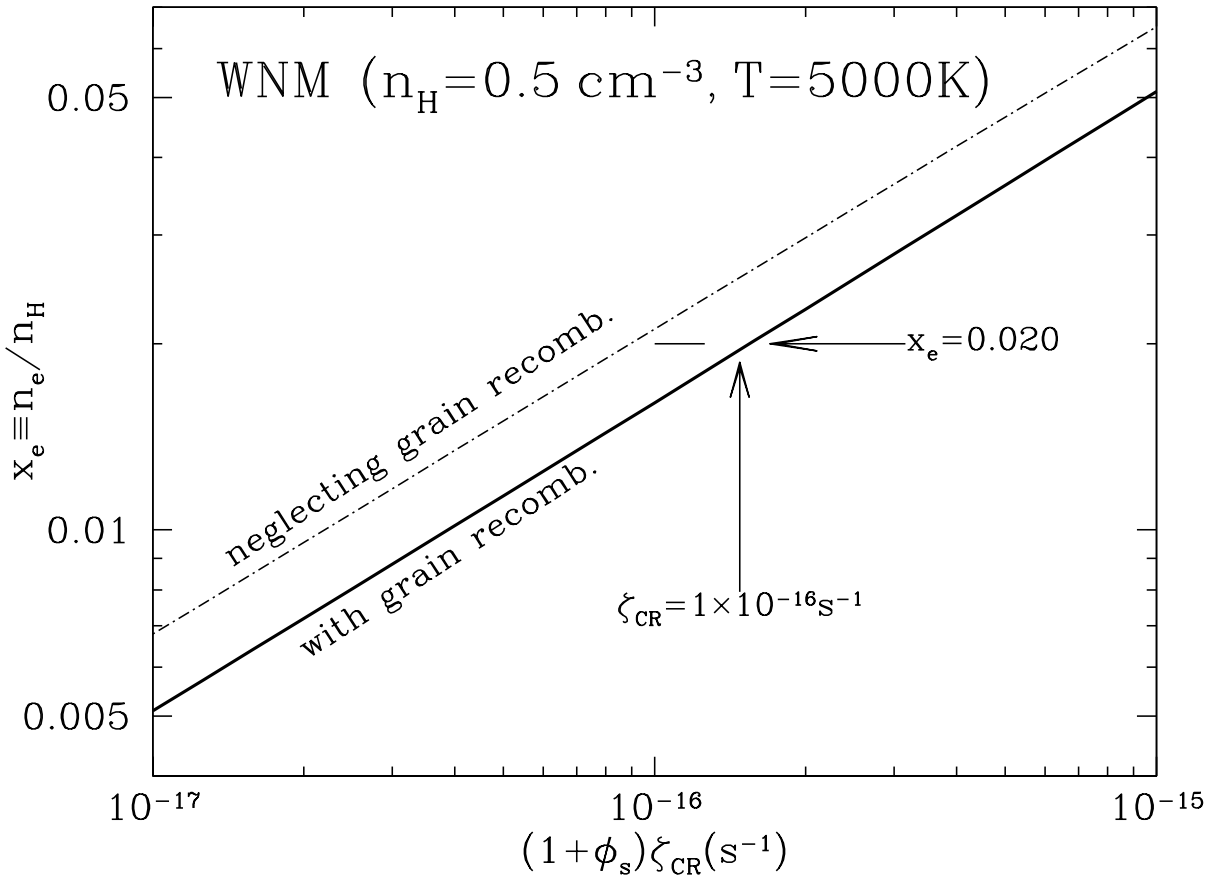


Figure 4.4: Draine 16.2

4.2.3 Warm H I Regions

Draine doesn't elaborate, but in the WNM, through similar arguments, the ionization fraction is a couple percent.

4.2.4 Diffuse Molecular Gas

In molecular environments, CRs and x-rays ionize H_2 to create H_2^+. The abundance of H_2^+ is rather low, but we can back out the cosmic ray ionization rate. If H_2^+ encounters a free electron, it will dissociatively recombine:

$$H_2^+ + e^- \rightarrow H + H\,. \tag{4.9}$$

Because the density of electrons is low, the following reaction is more likely:

$$H_2^+ + H_2 \rightarrow H_3^+ + H\,. \tag{4.10}$$

Various reactions remove H_3^+ from the gas. Draine lists 4, each with rate coefficients k_1, k_2, k_3, k_4.

As long as we can estimate the rate coefficients that determine how quickly H_3^+ is removed from the gas, we can determine the ionization rate. This is pretty powerful! Assuming x-ray ionization to be negligible, that means that we can more or less directly measure the cosmic ray ionization rate. Draine says that of our four processes, #1 and 2 are dominant.

We therefore have

$$\frac{n(H_3^+)}{n(H_2)} \approx \frac{2\zeta_{CR}(1+\phi_s)}{(k_1+k_2)n_H x_e} \tag{4.11}$$

(ϕ_s is the number of secondary ionizations for each primary) so

$$\zeta_{CR}(1+\phi_s) \simeq (k_1+k_2)n_H x_e \frac{n(H_3^+)}{2n(H_2)} \tag{4.12}$$

Direct measurements of H_3^+ have lead to values of $\zeta_{CR} \simeq 2 \times 10^{-16}$ s^{-1}.

4.2.5 Dense Molecular Clouds

If $A_V \gtrsim 3$, UV radiation cannot penetrate the cloud. C and S will be predominantly neutral in the gas phase. We do get ionization from CRs though, as usual.

Draine works out the ionization in dense molecular clouds. The fractional ionization is really low, $\sim 10^{-7}$ for $n_H \approx 10^4$ cm^{-3}.

4.3 Emission from a Plasma [Draine Chapter 10]

Now we have ions! How can we detect them? There are four main ways in which an astrophysical plasma emits:
(1) Free-free (Bremsstrahlung) continuum from free electrons in-elastically scattering off ions;
(2) Synchrotron continuum from relativistic electrons spirally around magnetic field lines;
(3) Free-bound continuum due to recombinations; and
(4) Bound-bound line emission from excited atoms (OK, I know this isn't a plasma, but as we'll see it's closely related)

4.3.1 Free-free emission (Bremsstrahlung)

Free-free emission is from electrons scattering off of ions. This scattering accelerates the electrons (changes their direction + velocity), and accelerated charges radiate. Ions are also accelerated, but much less due to their greater masses. We can ignore their contribution to the emission.

The electrons are unbound before and after the interaction, which leads to a continuous spectrum.

Free-free emission is the dominant continuum emission mechanism in thermal plasmas, and is therefore the dominant method in which plasmas cool, especially for high temperature (x-ray emitting) plasmas. The x-ray emission from clusters of galaxies is free-free. We won't get into this emission in this course on the ISM, but it is worth remembering that free-free emission is actually more important in the universe as a whole than it is in our Galaxy.

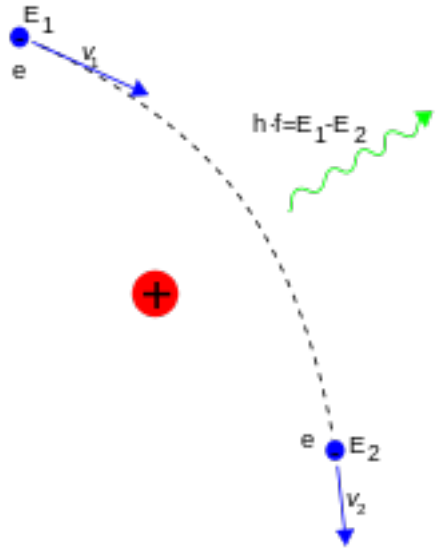


Figure 4.5: Bremsstrahlung emission from wikipedia: http://en.wikipedia.org/wiki/Bremsstrahlung

The derivation of Bremsstrahlung radiation is involved, and even Draine doesn't show it. The derivation on the Essential Radio Astronomy course is the best I have seen: http://www.cv.nrao.edu/course/astr534/FreeFreeEmission.html .

Draine gives the emissivity, the power per unit frequency, per unit volume, per steradian:

$$j_{\mathrm{ff},\nu} = \frac{8}{3}\left(\frac{2\pi}{3}\right)^{1/2} g_{\mathrm{ff,i}} \frac{e^6}{m_e^2 c^3}\left(\frac{m_e}{kT}\right)^{1/2} e^{-h\nu/kT} Z_i^2 n_e n_i \,, \tag{4.13}$$

where $g_{\mathrm{ff,i}}$ is the "Gaunt factor for free free transitions," and n_e and n_i are the electron and ion number densities, respectively. This can be simplified:

$$j_{\mathrm{ff},\nu} = 5.444 \times 10^{-41} g_{\mathrm{ff,i}} T_4^{-1/2} e^{-h\nu/kT} Z_i^2 n_e n_i \mathrm{erg\,cm^3\,s^{-1}\,sr^{-1}\,Hz^{-1}} \,. \tag{4.14}$$

Throughout this discussion, I will keep n_i and n_e as separate terms as Draine does. Remember that in general $n_e \approx n_i$, unless the plasma is really high temperature.

The Gaunt factor turns out to be very important. Classically, $g_{\mathrm{ff}} = 1$ (Kramers, 1923). In this case, the emissivity would be independent of frequency, for $\nu \ll kT/h$.

[Why does it depend on $n_e n_i$?] Two-body process.

The power is therefore

$$\Lambda_{\rm ff} = 4\pi \int_0^\infty j_{{\rm ff},\nu}\, d\nu = \frac{32\pi}{3}\left(\frac{2\pi}{3}\right)^{1/2} \frac{e^6}{m_e^2 hc^3}(m_e kT)^{1/2}\, \langle g_{\rm ff}\rangle_T\, Z_i^2 n_e n_i {\rm erg\,cm^3\,c^{-1}}\,, \qquad (4.15)$$

where

$$\langle g_{\rm ff}\rangle_T = \int_0^\infty \frac{dh\nu}{kT} e^{-h\nu/kT} g_{\rm ff}(\nu, T)\,. \qquad (4.16)$$

As we will see below, $\langle g_{\rm ff}\rangle_T$ is almost independent of frequency, so $\Lambda_{\rm ff} \propto n_e n_i T^{0.5}$.

4.3.2 The Gaunt Factor

From QM calculations, the Gaunt factor is

$$g_{\rm ff} \approx 6.155 (Z_i \nu_9)^{-0.118} T_4^{0.177}\,. \qquad (4.17)$$

This is only valid for frequencies between the plasma frequency and kT/h.

We see that the frequency dependence of the Gaunt factor is really shallow. Draine mentions that this approximation is good to $\pm 10\%$ for $0.14 < Z_i\nu_9/T_4^{3/2} < 250$.

This leads to

$$j_{{\rm ff},\nu} = 3.35 \times 10^{-40} Z_i^{1.882} n_e n_i \nu_9^{-0.118} T_4^{-0.323} {\rm erg\,cm^3\,s^{-1}\,sr^{-1}\,Hz^{-1}} \qquad (4.18)$$

within the frequency range specified above, $\nu_p \ll \nu \ll kT/h$.

The free-free emissivity in the radio and microwave goes as $\nu^{-0.118}$!

Draine also gives the frequency averaged Gaunt factor, which leads to

$$\Lambda_{\rm ff}(T) \approx 1.422 \times 10^{-25} \left\{1 + \frac{0.44}{1 + 0.058[\ln(T/10^{5.4} Z_i^2\ {\rm K})]^2}\right\} T_4^{1/2} Z_i^2 n_i n_e \frac{\rm erg\,cm^3}{\rm s}\,, \qquad (4.19)$$

Keep this in mind for later.

4.3.3 Opacity and Optical Depth

We know that in LTE, the energy levels are populated according to a thermal distribution, and we can apply Kirchoff's law: $\kappa_\nu = j_\nu / B_\nu(T)$. This again is just saying that there must be equilibrium between emission and absorption, otherwise the temperature would be changing. Therefore,

$$\kappa_{{\rm ff},\nu} = \frac{4}{3}\left(\frac{2\pi}{3}\right)^{1/2} \frac{e^6}{m_e^{3/2}(kT)^{1/2} hc\nu^3} \left[1 - e^{-h\nu/kT}\right] Z_i^2 n_i n_e g_{ff}\,. \qquad (4.20)$$

In the Rayleigh-Jeans radio limit, we can expand the exponential and input our analytical fit to the Gaunt factor to get

$$\kappa_{\mathrm{ff},\nu} \approx 1.091 \times 10^{-25} Z_i^{1.882} T_4^{-1.323} \nu_9^{-2.118} n_i n_e \,\mathrm{cm}^{-1} \,. \tag{4.21}$$

We see that free-free absorption becomes strong at low frequencies. In fact, there is a low frequency cutoff to the Bremsstrahlung radiation caused by free-free absorption.

We can of course also define the optical depth [following essential radio astronomy course]. Worrying only about proportionalities,

$$\tau_{\mathrm{ff},\nu} = \int \kappa_{\mathrm{ff},\nu}\, ds \propto Z_i^{1.882} T_4^{-1.323} \nu_9^{-2.118} n_i n_e \tag{4.22}$$

[Essential radio astronomy course] We can now define two regimes: optically thick and optically thin. In the radio regime, in the optically thick limit, $\tau \gg 1$ and the flux density

$$S_\nu \approx B_\nu(T) = \frac{2kT\nu^2}{c^2} \propto \nu^2 \,. \tag{4.23}$$

In the optically thin limit, $\tau \ll 1$, and the flux density

$$S_\nu \approx B_\nu(T) = \frac{2kT\nu^2}{c^2}\tau_{\mathrm{ff}} \propto \nu^2 \nu^{-2.118} \propto \nu^{-0.12} \,. \tag{4.24}$$

This gives rise to the following plot.

Where these two regime intersect, at $\tau = 1$, is the "turnover frequency." The turnover frequency can be as high as 10^{10} GHz, but 1 GHz is more common.

In the optically thin limit, the intensity detected in observations of free-free emission are linearly proportional to the optical depth. In the radio regime, $I_\nu = B_\nu(T_e)\tau$ and $\tau_{\mathrm{ff},\nu} \propto \int n_i n_e \, ds$. For convenience we can therefore define a parameter called the "emission measure," EM:

$$\mathrm{EM} = \int n_e^2 \, ds \,\mathrm{cm}^{-6}\,\mathrm{pc} \,. \tag{4.25}$$

[Draine here of course defines EM differently from the rest of the world, and uses $n_e n_i$. Typical Draine!] To be useful we must assume that $n_e \approx n_i$. Notice the units! The emission measure for an H II region 1 pc in diameter, with a density of 1 cm^{-3} is 1 cm^{-6} pc.

Therefore, $I_\nu \propto T^{-1.323}\nu^{-0.118} EM$ for optically thin emission, and $I_\nu \propto T\nu^2$ for optically thick emission.

4.3.4 Synchrotron Radiation

Cosmic rays are astrophysical particles (electrons, protons, and heavier nuclei) with extremely high energies. Cosmic-ray electrons in the Galactic magnetic field emit the synchrotron radiation that accounts for most of the continuum emission from our Galaxy at

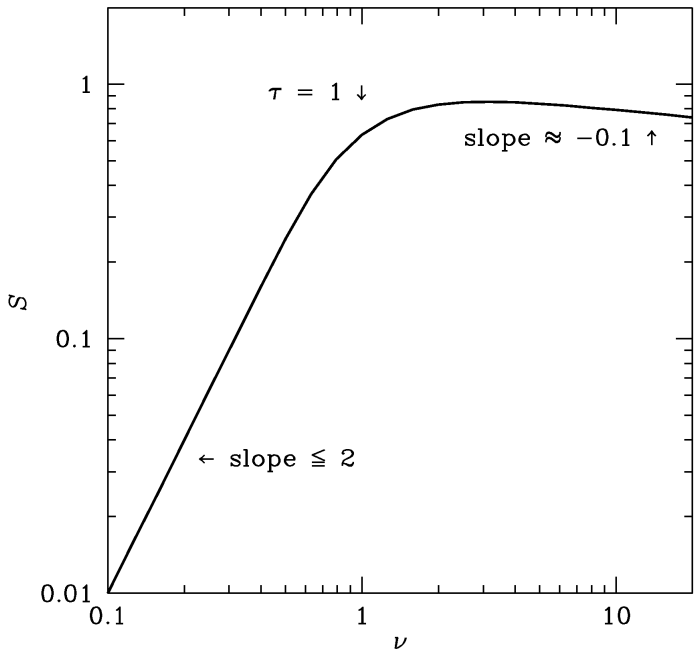


Figure 4.6: Bremsstrahlung spectrum from Essential Radio Astronomy Course.

frequencies below about 30 GHz. They spiral around the magnetic field lines. In the same way that Bremsstrahlung electrons produce free-free. At low radio frequencies, synchrotron dominates the radiated power of the Galaxy. Synchrotron is non-thermal emission, because it does not depend on the temperature of the electrons.

[What emits synchrotron?] Pulsars, SNRs, AGN, radio galaxies, etc.

Synchrotron radiation is polarized, because of the preferred axis along the magnetic field lines. The polarization axis is perpendicular to the magnetic field direction as projected onto the sky.

The individual electrons each radiate power

$$P = \frac{2}{3}\frac{q^2 a^2}{c^3} \text{ (cgs units)}, \tag{4.26}$$

where a is the acceleration, q is the charge, and c is the speed of light. This is Larmor's formula.

The observed synchrotron spectrum is the summation of the contributions from individual electrons. It has been found that the distribution of electron energies is a power law: $n(E) \propto E^{-p}$. Obviously, there are more slow moving electrons than fast moving ones. A typical value for p in the ISM is 2.4.

When we average over all electrons, in the optically thin limit we find

$$\frac{dP}{dV\,d\nu} \propto n_e B^{(p+1)/2}\,\nu^{-(p-1)/2},. \tag{4.27}$$

We can see that the power increases with increasing electron density and magnetic field, but decreases with increasing frequency.

The synchrotron spectrum has a high frequency cutoff when we run out of high energy electrons. Like Bremsstrahlung, there is a low frequency cutoff when absorption becomes dominant. The absorption in this case is synchrotron self-absorption.

The emissivity j for synchrotron is just the power from above divided by 4π. For free-free, we could assume that the source function was just $S_\nu = j_\nu/\kappa_\nu = B_\nu(T)$, and we could easily find κ_ν. For non-thermal emission, we can't assume that any longer. I will state without proof that

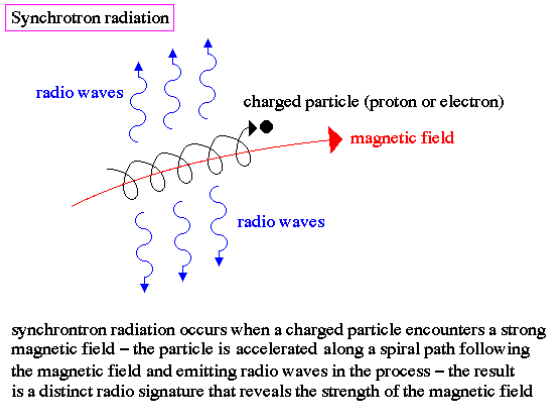
$$\kappa_\nu \propto n_e B^{(p+2)/2}\nu^{-(p+4)/2} \tag{4.28}$$

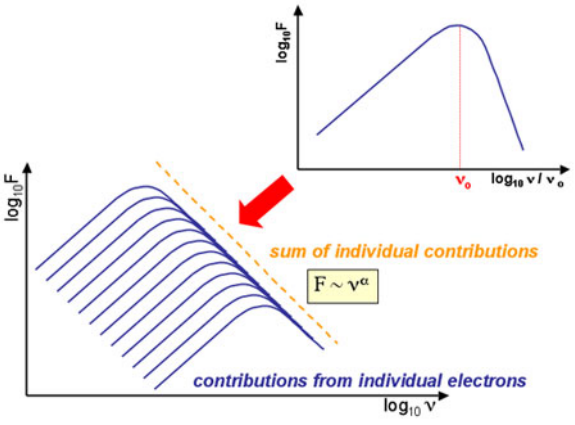
and therefore

$$S_\nu = \frac{j_\nu}{\kappa_\nu} \propto B^{-1/2}\nu^{5/2}\,. \tag{4.29}$$

Note that the source function is independent of the power law index of the relativistic electrons. At low frequencies, the synchrotron radiation is optically thick, and $I_\nu \propto S_\nu \propto B^{-1/2}\nu^{5/2}$. At high frequencies, the intensity will be proportional to the emissivity coefficient: $I_\nu \propto j_\nu \propto \nu^{-(p-1)/2} \propto \nu^{-0.7}$.

The frequency at which the spectrum transitions depends on the depth of the region, the relativistic electron number density, and magnetic field strength.





One last point: the spectral index of synchrotron radiation is the power-law index of the flux density distribution:

$$S \propto \nu^\alpha\,. \tag{4.30}$$

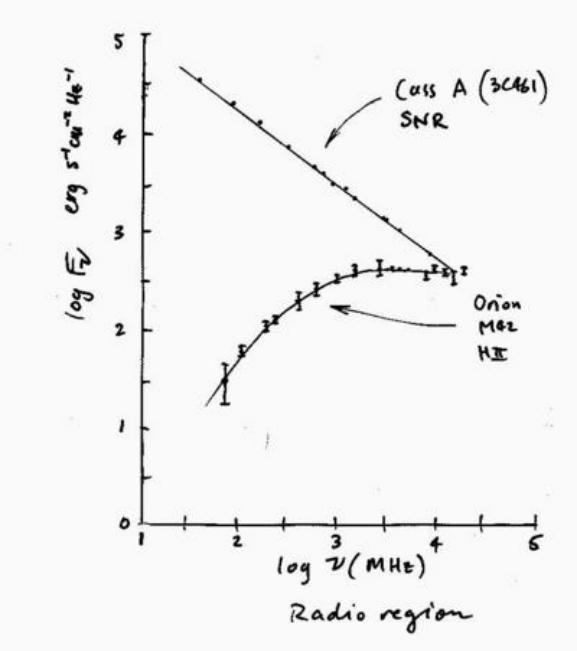
Typical values are:
-0.7 for radio galaxy

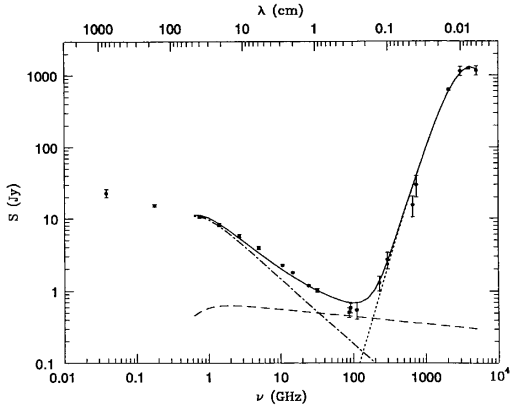
-0.5 for SNRs
-3 to -2 for pulsar
-1 to 1 for AGN.

Contrast this with thermal (Bremsstrahlung) emission, which has a spectral index of $\alpha \simeq 0.1$ for optically thin thermal emission. This allows us to distinguish between thermal and non-thermal emission.

Some people define α as positive in the above equation and some as negative. Be careful!!!

4.3.5 All together (the figures)





4.3.6 Free-bound Transitions

Also known as "radiative recombination."

$$j_{\mathrm{fb},\nu} = \frac{g_b}{g_e g_i} \frac{h^4 \nu^3}{(2\pi m_e kT)^{3/2} c^2} e^{(I_b - h\nu)/kT} \sigma_{\mathrm{b,pi}}(\nu) n_e n_i \,, \tag{4.31}$$

where g_b is the degeneracy of bound state b, $g_e = 2$ is the degeneracy of the free electron, g_i is the degeneracy of the ion, and I_b is the energy required to ionize from bound state b.

Each bound state b contributes continuum beginning at $h\nu = I_b$, cut off at high frequencies by the factor $e^{(I_b - h\nu)/kT}$.

For hydrogen, $I_b = I_{\mathrm{H}}/n^2$ and $g_b/g_i = 2n^2$.

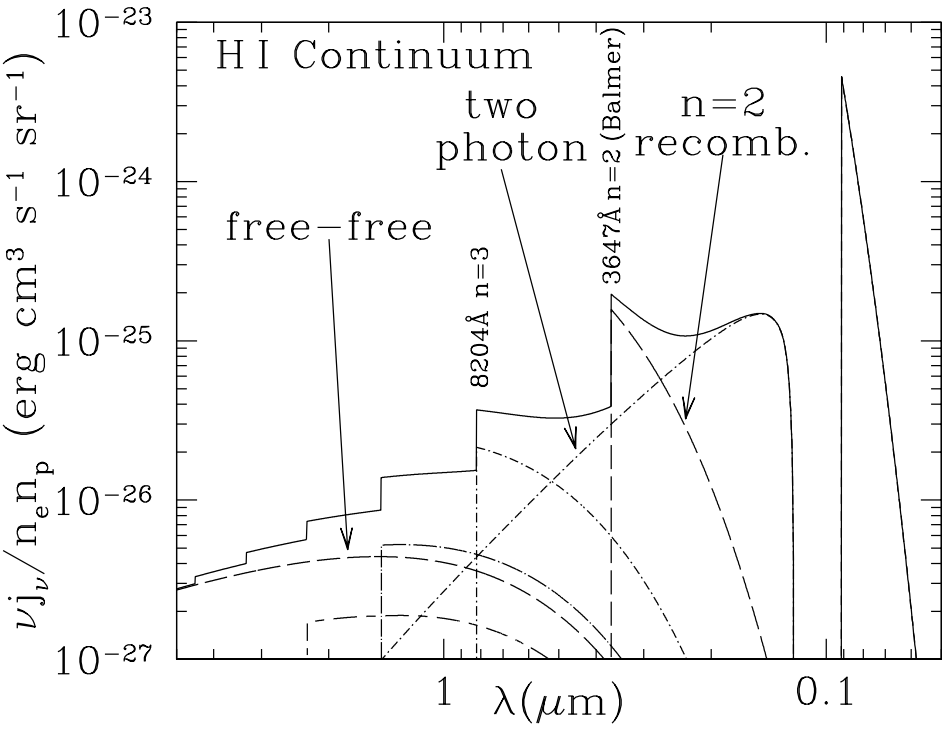


Figure 4.7: The emission spectrum of a $T = 8000\,\mathrm{K}$ hydrogen plasma.[Draine 10.2]

4.4 Bound-Bound Transitions

Bound-bound transitions can result from excitation (from collisions or radiation), or from recombination. We'll deal with these in turn. Because of the two excitation mechanisms and the different physics of emission, this subsection is by far the most involved.

For excitation, we have already sketched out the relevant equations, but it's worth repeating here. Why is this section in the "ionized gas" portion of the course? Almost all the bound-bound transitions we care about are from ions (or newly recombined neutrals).

4.4.1 Quantum Mechanics Review [Draine chapter 4]

We first have to get our notation down. For electronic transitions, for atomic or ionized gas, the goal of the notation is to specify the energy levels completely. Our discussion here pertains to ionization, because we are almost interested in the energy levels of ionized atoms. Everything is also applicable to neutral atoms as well.

Spectroscopic Notation

Ionized Gas Each electron that gets ejected gets expressed with an additional roman numeral. H I is neutral hydrogen, H II is ionized hydrogen, OIII is doubly ionized oxygen, etc.

Notation for atoms We can approximate electrons in atoms as having "single electron" orbitals. These orbitals are characterized by the quantum numbers:

n: the principle quantum number. Classically, this is the *size* of the orbital. Takes integer values of 1,2,3,...

ℓ: the orbital angular momentum in units of $\hbar$. This is the *shape* of the orbital. Takes values $\ell = 0, 1, 2, ..., n-1$, for n values. The values of ℓ are given additional nomenclature: $\ell = 0, 1, 2, 3$ are referred to as s, p, d, f ("sharp," "principle," "diffuse," and "fundamental" were used originally for hydrogen to describe the spectra. Examples are $ns \rightarrow np$, $np \rightarrow ns$, $nd \rightarrow np$, $nf \rightarrow nd$).

m_ℓ: the projection of the angular momentum/$\hbar$ onto the z axis. We see in the figure that there is only one way to orient a sphere ($\ell = 0$), but multiple ways for the other orbital shapes. m_ℓ (sometimes called m_z, as in Draine) takes values $-\ell ... -1, 0, 1, \ell$, for a total of $2\ell + 1$ values. If there is no external magnetic field, while the quantum numbers are still present, the energies are degenerate.

s: the spin quantum number. Has value of 1/2 for electrons.

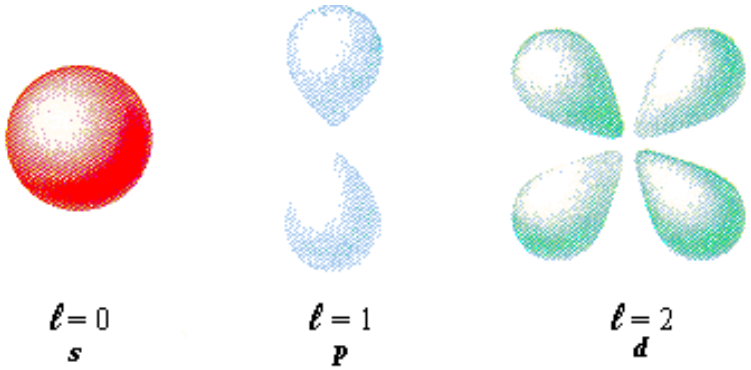
m_s: The projection of the electron spin angular momentum onto the z axis. Takes values of $-1/2, +1/2$. Again, if there is no applied field, the two states are degenerate.

$j = |\ell \pm s|$: the total angular momentum. Both ℓ and s are angular momentum. The use of j is called spin-orbit (or $L - S$) coupling. This gives rise to "fine structure" lines discussed below.

m_j: projection of the total angular momentum along a specified axis. $m_j = m_\ell + m_s$ and $|m_\ell + m_s| \leq j$.

In the simplest case of a single electron, there are $2(2\ell+1)$ distinct electronic wave functions for each $n\ell$ pair (called a "sub-shell"). The $2\ell + 1$ term is from the m_ℓ quantum number,

and the factor of 2 is from the spin.



Pauli Exclusion Principle For multiple electron atoms, we cannot have the same wavefunction (set of quantum numbers) for any two electrons. This gives us the limit for any sub-shell of $2(2\ell+1)$, 2 for s, 6 for p, 10 for d, etc..

The sub-shells increase in energy as: $n\ell = 1s^2, 2s^2, 2p^6, 3s^2, 3p^6, 4s^2, 3d^{10}, 4p^6$.... We see that energies increase with increasing n and ℓ, but since the energies associated with n and ℓ are roughly comparable, the ordering of sub-shell energies is more complicated. This gives rise to two electrons in the first n shell, 8 in the second, 18 in the third, and 32 in the fourth. This can easily be seen in the periodic table.

We denote the shell occupancy with $n\ell^{\#elec}$, so for Na with 11 electrons, the ground state configuration is $1s^2 2s^2 2p^6 3s$ (the "1" on the $3s$ term is implied). Often, we would just shorten this to call it a "$3s$" atom.

Multiple Electrons When multiple electrons are involved, single electron QNs are not as useful. We need a method for specifying all the electrons, with MOAR QUANTUM NUMBERS!!
L: the total angular momentum of all the electrons.
S: the total spin angular momentum of all the electrons.
$J = |L+S|$: the total angular momentum. Like its little brother j, J ranges from $L-S$ to $L+S$. It need not be an integer, since S can be $1/2$.
Using the J QN for L and S is called the "L-S coupling" approximation.

L-S coupling is just the coupling between the L and S vectors to produce the new quantum number J. Valid when external magnetic field weak, atomic number $\lesssim$ 40. This is an approximation of sorts, albeit a very good one.

All Together We now have all that we need to fully specify the energy of an atom. The full spectroscopic notation is:

$$^{2S+1}\mathcal{L}^p_J\,, \tag{4.32}$$

where $\mathcal{L} = S, P, D, F$ for $L = 0, 1, 2, 3$, and p is the "parity". L takes all possible values of $\sum \ell$, and so there are multiple different possible spectroscopic notations for atoms with

multiple electrons in an unfilled subshell. Parity can be "even" or "odd". The parity must change in electric dipole transitions (see below), and so is an important quantity to note. If the quantity $\Sigma_i \ell_i$ is even, the parity is even, and the same if it is odd.

This notation gives us all the information we need (usually). It does not include the splitting of degeneracy in magnetic fields, or the hyperfine-structure splitting that gives rise to the H 21 cm line.

Each LS combination is called a "term": ${}^{2S+1}\mathcal{L}$. There may be multiple values of J for each term. If so, these are called "levels." The levels tells you how L and S are coupled. The "multiplicity" of a term with total spin S and total orbital angular momentum L is $g = (2S+1)(2L+1)$. $2S+1$ is the "spin multiplicity."

Filled orbitals have no net spin or orbital angular momentum, so we express the state of atoms only using their valence electrons.

Let's do an example. The ^{3}P term has $L = 1$ and $S = 1$. Adding L and S vectorially, we get $J = 0, 1, 2$. The ^{3}P term therefore has three possible energy configurations: ${}^3\mathrm{P}_0$, ${}^3\mathrm{P}_1$, and ${}^3\mathrm{P}_2$. This is called "fine-structure" splitting. Before, we had "hyperfine" structure for H I. Terms like ^{3}P with three fine-structure values are called "triplets," those with one "singlets," etc. Those with multiple terms are collectively called "multiplets."

Let's take a look at Draine Figure 4.1.

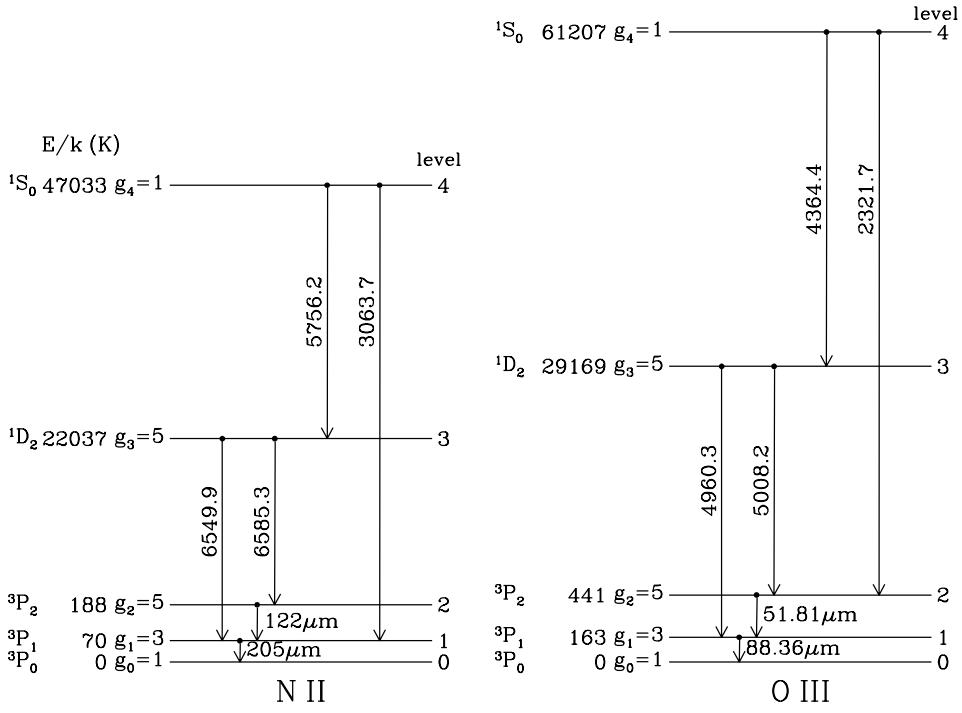


Figure 4.8: Draine Figure 4.1 showing energy levels of OIII and NII.

Both NII and OIII have six electrons, arranged as $1s^2 2s^2 2p^2$. We are concerned only with the $2p^2$ electrons in the un-filled subshell.
This gives rise to three terms: ^{1}D, ^{3}P, and ^{1}S (derivation of these three terms is slightly involved; see footnote 1 of Draine Chapter 4. Basically, you determine all the terms, then you have to apply the Pauli Exclusion Principle).
The ^{1}D and ^{1}S terms must be singlets, since $S = 0$, so we have $^1\mathrm{D}_2$ and $^1\mathrm{S}_0$. The ^{3}P term is a triplet as we saw earlier, so we have $^3\mathrm{P}_0$, $^3\mathrm{P}_1$, and $^3\mathrm{P}_2$.

The energy level diagram shows how the energies are arranged. [Kwok pg 83] Hund's empirical rules for atomic energies says that:
Higher S = lower energy
Higher L = lower energy
Higher J = higher energy if the shell is less than half-filled, lower energy if more than half filled
If exactly half-filled, lowest energy term is always S ($L = 0$), but other terms do not obey above rules. We get higher energies with: higher J, and usually higher energies with the lowest possible values of L and S, so $^1\mathrm{S}_0$ is higher energy than $^1\mathrm{D}_2$.

Selection Rules

Electric Dipole Radiation has the following selection rules. Some of these are ruled out by others, so you should think of each one as being independent.
(1) Parity must change.
(2) $\Delta L = 0, \pm 1$, but $\Delta L = 0 \rightarrow 0$ forbidden. ($\Delta L = 0$ ruled out by rule # 1).
(3) $\Delta J = 0, \pm 1$, but $J = 0 \rightarrow 0$ forbidden.
(4) Only one single electron wavefunction changes, with $\Delta \ell = \pm 1$. This one took me a while initially to wrap my head around. There are many ways that you could have $\Delta \ell = 0, \pm 1$ if you allow for multiple electrons to transition. The electron carries away angular momentum, and this angular momentum must come from the atom; thus the rule.
(5) $\Delta S = 0$ (Spin does not change). This is another way of saying that the transitions must be between states with the same multiplicity.

Electric dipole transitions are of course the strongest. They have the largest Einstein A coefficients, which leads to the smallest lifetimes in their excited states. They are denoted as, e.g., NII+wavelength.

In the ISM, the density is low. Really low. All atoms have excited states for which there is no dipole radiation possible. In laboratory densities, these states are collisionally de-excited. In the ISM, this is not the case, and we can get magnetic dipole and electric quadrupole radiation out (magnetic quadrupole, and electric and magnetic octopole radiation are possible but not observed in the ISM to my knowledge).

Transitions that satisfy electric dipole selection rules 1 to 4, but fail rule #5 are called "semi-forbidden," "intercombination," or "intersystem" transitions. These are denoted, e.g.,

NII]+wavelength. Transitions are $\sim 10^6$ times weaker than electric dipole radiation.

Transitions that fail at least one of the other four selection rules are called "forbidden." While these transitions are not strictly forbidden, we never see them in the lab (only in the low density ISM). Electric quadrupole and magnetic dipole transitions are both forbidden. We denote forbidden transitions with, e.g. [NII]+wavelength. They are $\sim 10^2 - 10^6$ times weaker than permitted transitions.

Hydrogen [Draine 9.7] It is worth spending a little time on ionized hydrogen, the most abundant ionized specites. Hydrogen of course has a single electron, so its state is $1s$. From this state it can transition to the $2p, 3p, 4p...$ states. Transitions into or out of the $n = 1$ level are called Lyman, $n = 2$ Balmer, $n = 3$ Brackett, $n = 4$ Paschen, $n = 5$, Pfund, and $n = 6$ Humpfreys (page 483).

Because of how the hydrogen energy levels are arranged, the Lyman series starts in the UV, the Balmer series in the visible, and the others in the IR.

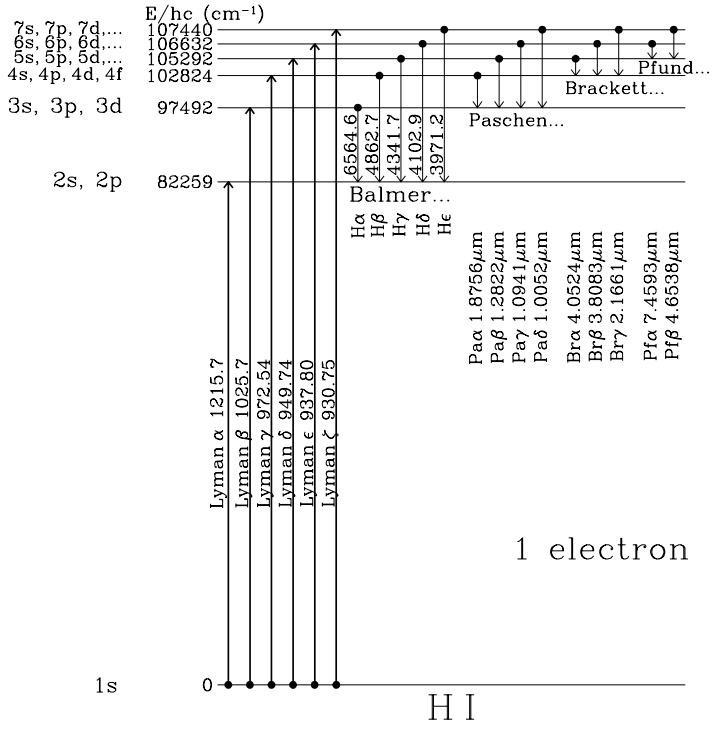


Figure 4.9: Draine Figure E.1 showing energy levels of H.

Transitions between adjacent levels are called α-transitions, $\Delta n = 2$ are β, etc. Lyman α is the transition from $n = 2 \rightarrow 1$. Balmer α, i.e. Hα is $n = 3 \rightarrow 2$.

The $2p$ state is actually a doublet. $s = 1/2$ and $\ell = 1$, so $g = (2s+1)(2\ell+1) = 2$. This leads to two Lyman α lines (here in absorption): $^2\mathrm{S}_{1/2} \rightarrow {}^2\mathrm{P}^o_{1/2}$ and $^2\mathrm{S}_{1/2} \rightarrow {}^2\mathrm{P}^o_{3/2}$. The energy level difference between these transitions is minimal compared to the intrinsic line-width. We can therefore treat these energies as being degenerate.

The "Lyman limit" is at 13.6 eV. Photons with energies greater than 13.6 eV (91.16 nm) are quickly absorbed in the ISM by hydrogen. These photons basically do not exist in the diffuse ISM.

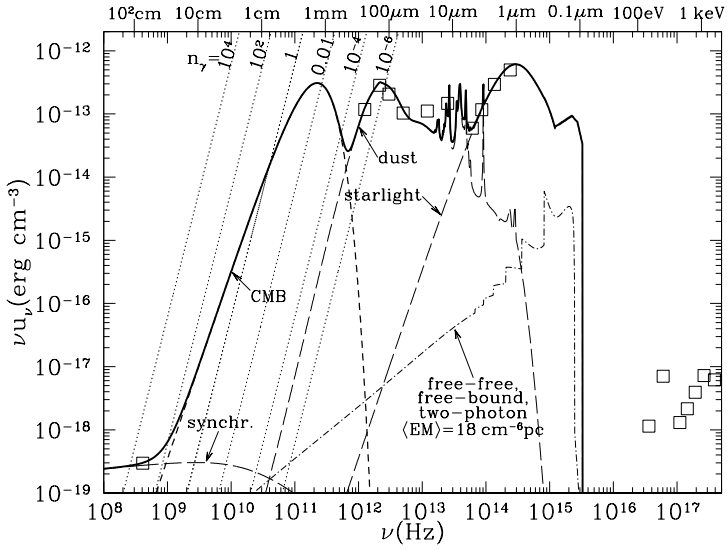


Figure 4.10: Draine Figure 12.1 showing hydrogen energy levels.

Line Strengths [Kwok 5.2] Line strengths depend on the population in the state and the Einstein A coefficients. In the case of lines within a multiplet, the Einstein As will be similar, and the line-strengths will be determined by the population in the state:
The sum of the strengths of all lines of a multiplet that end on a common final level is proportional to the weight $2J + 1$ of the level.
The sum of the strengths of all lines of a multiplet that start from a common initial level is proportional to the weight $2J + 1$ for the initial level.

The reason for these rules is simple: the degeneracy of the level is $2J + 1$. Remember previously that the degeneracy for H I is $g_\mathrm{u} = 3$ for the upper state, and $g_\ell = 1$ for the lower state. This implied that 3/4 of all H I is in the upper (spin aligned) state, and 1/4 in the lower (spin anti-aligned) state. The degeneracy for H I is $g = 2S + 1$, for hyperfine levels, so $2J + 1$ for fine-structure should be expected.

We see these degeneracy values in the energy level diagrams.

For example, transitions between levels of the ^4P term and the ^4D term. The ^4D term ($S = 3/2, L = 2$) can take values of $J = 7/2, 5/2, 3/2, 1/2$. The ^4P term ($S = 3/2, L = 1$) can take values $J = 5/2, 3/2, 1/2$.

The highest J states have the highest degeneracy, so there should be more transitions into and out of these states. The degeneracy for the $7/2, 5/2, 3/2, 1/2$ states are $16, 12, 8, 4$. The strongest line will be the $J = 7/2 \to 5/2$ line, and the weakest will be $J = 1/2 \to 1/2$.

4.4.2 Excitation

Radiative excitation is relatively straightforward, so let's deal with collisions first. The two-body collision rate coefficient is

$$\langle \sigma v \rangle_{AB} \equiv \int_0^\infty \sigma_{AB}(v) v f_v dv \,, \tag{4.33}$$

where A and B are two species. This term has units of $\mathrm{cm}^3\,\mathrm{s}^{-1}$. We can usually assume a MB distribution for f_v. The quantity $\langle \sigma v \rangle$ is sometimes called the "excitation rate coefficient" $\gamma_{\ell u}$, which tells you the number of excitations from state ℓ to state u in units of $\mathrm{cm}^{-3}\ \mathrm{s}^{-1}$. Therefore for collisions with electrons (almost always assumed), $n_e n_\ell \gamma_{\ell u}$ gives the number of excitations per cc per second.

As we did before, we can assume that the upwards and downwards collisions balance ("detailed balance"). Therefore, for collisions with electrons,

$$n_\ell n_e \gamma_{\ell u}(T) = n_u [A_{u\ell} + n_e \gamma_{u\ell}(T)] \,, \tag{4.34}$$

which reduces to

$$\frac{n_u}{n_\ell} = \frac{\gamma_{\ell u}}{\gamma_{u\ell}} \left(1 + \frac{A_{u\ell}}{n_e \gamma_{u\ell}} \right)^{-1} = \frac{g_u}{g_l} e^{\Delta E / k T_{\mathrm{ex}}}. \tag{4.35}$$

This equation has high and low density regimes, as we saw before. At high densities, $A_{u\ell}$ is unimportant. At low densities, the collision rate is unimportant.

We usually further parametrize σ in terms of the "collisional strength" Ω, such that

$$\sigma_{\ell u}(v) = \frac{\pi \hbar^2}{m_e^2 v^2} \frac{\Omega_{\ell u}}{g_\ell} \,, \tag{4.36}$$

(this is sometimes written as a function of energy instead of velocity) or for MB speed distribution

$$\langle \sigma_{\ell u} v \rangle = \gamma_{lu} = \frac{h^2}{(2\pi m_e)^{3/2}} \frac{1}{(kT)^{1/2}} \frac{\Omega_{\ell u}}{g_\ell} = \frac{8.629 \times 10^{-8}}{\sqrt{T_4}} \frac{\Omega_{\ell u}}{g_\ell} \,. \tag{4.37}$$

The collision strength is approximately unity for most transitions, and is almost independent of temperature. Draine tabulates some useful values in Appendix F. Keep in mind that $\Omega_{u\ell} = \Omega_{\ell u}$ in LTE.

4.4.3 (Collisionally Excited) Bound-bound Emission Lines [Kwok 5.3]

Various fine-structure lines are important for our understanding of the ISM. Although forbidden, these lines have high Einstein A's due to the ν^3 dependence, and so emit readily.
[CI] $^3\mathrm{P}_2 \rightarrow {}^3\mathrm{P}_1$ at $360\,\mu$m and $^3\mathrm{P}_1 \rightarrow {}^3\mathrm{P}_0$ at $609\,\mu$m. Carbon is abundant and these transitions are important in low-energy environments.
[OIII] $^3\mathrm{P}_2 \rightarrow {}^3\mathrm{P}_1$ at $52\,\mu$m and $^3\mathrm{P}_1 \rightarrow {}^3\mathrm{P}_0$ at $88\,\mu$m. Ditto, but for harder radiation fields.
[NII] $^3\mathrm{P}_2 \rightarrow {}^3\mathrm{P}_1$ at $122\,\mu$m and $^3\mathrm{P}_1 \rightarrow {}^3\mathrm{P}_0$ at $205\,\mu$m. Note that CI, OIII, and NII have the same electronic configurations.
[CII] $^2\mathrm{P}_{3/2} \rightarrow {}^2\mathrm{P}_{1/2}$ at $158\,\mu$m. Because of carbon's lower ionization potential compared to hydrogen, there are plenty of photons around to ionize C. The [CII] lines are really bright.

These lines are incredibly important for the energy balance of the ISM. The [CII] line is the most efficient cooling line of the ISM in fact!

Total luminosities measured by *COBE FIRAS*.

Species	Transition	λ (μm)	$\log L(L_\odot)$
CO	$J = 2 \rightarrow 1$	1302	4.9
CO	$J = 3 \rightarrow 2$	867.2	5.1
CO	$J = 4 \rightarrow 3$	650.4	4.1
[CI]	$^3P_1 \rightarrow ^3\,P_0$	609.1	5.3
CO	$J = 5 \rightarrow 4$	519.8	5.0
[CI]	$^3P_2 \rightarrow ^3\,P_1$	370.4	5.5
[NII]	$^3P_1 \rightarrow ^3\,P_0$	205.3	6.7
[CII]	$^2P_{3/2} \rightarrow ^2\,P_{1/2}$	157.7	7.7
[NII]	$^3P_2 \rightarrow ^3\,P_1$	121.9	6.9

4.4.4 Recombination [Draine Chapter 14]

We talked about ionization: radiative, collisional, and by cosmic rays. The second half of the equation of course is recombination. How are ions and electrons able to recombine?
Radiative recombination: $X^+ + e^- \rightarrow X + h\nu$.
Dielectric recombination: $X^+ + e^- \rightarrow X^{**} \rightarrow X + h\nu$
Three-body recombination: $X^+ + e^- + e^- \rightarrow X + e^-$
Charge exchange: $X^+ + Y \rightarrow X + Y^+$
Dissociative recombination: $AB^+ + e^- \rightarrow A + B$
Neutralization by grain: $X^+ + \mathrm{grain} \rightarrow X + \mathrm{grain}^+$
Draine mentions that the relative importance of these depends on the ion in question and the physical conditions of the plasma. We will only cover the most important recombination reactions for most astrophysical plasmas, namely radiative and dielectric recombination.

Radiative Recombination

The most obvious recombination reaction is radiative recombination, where an ion and an electron recombine to create a neutral atom, and a photon is released corresponding to the

difference in energy between the free and bound states (this is free-bound radiation, to be covered next lecture).

The thermal rate coefficient for recombination into level $n\ell$, $\alpha_{n\ell}$ in units of $\mathrm{cm}^3\,\mathrm{s}^{-1}$, is exactly the same as that from collisional ionization, aside from the integration limits and different cross section subscripts:

$$\alpha_{n\ell} = \int_0^\infty \sigma_{\mathrm{rr},n\ell}(E) v f_E dE\,, \tag{4.38}$$

which for a Maxwellian distribution is

$$\alpha_{n\ell} = \left(\frac{8kT}{\pi m_e}\right)^{1/2} \int_0^\infty \sigma_{\mathrm{rr},n\ell}(E) \frac{E}{kT} e^{-E/kT} \frac{dE}{kT}\,. \tag{4.39}$$

The free-bound emission has an energy of $h\nu = I_{n\ell} + E$, where $I_{n\ell}$ is the energy required for ionization from level $n\ell$. For hydrogen, for example, $I_{n\ell} = I_H/n^2$, the binding energy of the electron in level $n\ell$.

Sometimes authors speak of a recombination rate in $\mathrm{cm}^{-3}\ \mathrm{s}^{-1}$:

$$N_R = n_{\mathrm{ion}} n_e \alpha \approx n_e^2 \alpha \tag{4.40}$$

Radiative Recombination for Hydrogen

Being the most abundant element by far, hydrogen's recombination rates will have the largest impact on the ISM. Draine's Table 14.1 lists radiative recombination coefficients for hydrogen. Remember, the energies are almost degenerate, but the different ℓ states are real and have different recombination rate coefficients.

Recombinations directly to level $n = 1$ result in photons of energy I_H that can ionize hydrogen. This reaction therefore does not alter the plasma, since the emitted photon can go on to ionize another atom. This process is efficient in dense environments, and recombinations down to level $n = 1$ are quickly absorbed near the point of emission. Aside from transport of the ionization energy a short distance, this has virtually no effect on the ionization state of the gas. In low density environments, this effect is less important.

We can therefore create two recombination coefficients for the two different environments (Baker & Menzel 1938):
Case A: Optically thin to ionizing radiation; every photon emitted during recombination escapes. Sum over all $n\ell$ to get α_A (Table 14.1).
Case B: Optically thick to radiation just above $I_H = 13.60\,\mathrm{eV}$; radiation from recombinations directly to $n = 1$ are quickly absorbed. Sum over all $n\ell$ $n \geq 2$ to get α_B.

The Case A and Case B coefficients are just the summation over all recombination coefficients. So Case A is $\sum_{i=1}^{\infty} \alpha_i$ and Case B is $\sum_{i=2}^{\infty} \alpha_i$, where i represents $n\ell$.

These two cases have application in two regimes found in the ISM. Case B recombination is appropriate in H II regions, where the density of H is large enough (although still only $n \gtrsim 1\ \mathrm{cm}^{-3}$).

The HIM with $T \gtrsim 10^6\,\mathrm{K}$ is collisionally ionized and has little neutral hydrogen. Because of the high temperature, the free-bound emission is hard. Case A is an excellent approximation here. For the more diffuse gas of the WIM, either may be appropriate, depending on the density.

We can approximate the Case A and B recombination rate coefficients for hydrogenic elements:

$$\alpha_A(T) \approx 4.13 \times 10^{-13} Z (T_4/Z^2)^{-0.7131-0.0115\ln(T_4/Z^2)} \tag{4.41}$$

and

$$\alpha_B(T) \approx 2.54 \times 10^{-13} Z (T_4/Z^2)^{-0.8163-0.0208\ln(T_4/Z^2)} \tag{4.42}$$

T_4 here is the "electron temperature" of the plasma. In a thermal plasma, the electrons and ions will have roughly the same energies, but due to the mass differences, different velocities and temperatures.

These approximations are different from what I learned in grad school. Dyson & Williams advocate for

$$\alpha_B = 2 \times 10^{-10} T_e^{-3/4} = 2 \times 10^{-13} T_4^{-3/4}\,. \tag{4.43}$$

These are slightly different at $T = 10^4\,\mathrm{K}$, as the Draine approximation gives 2.54×10^{-13} and the Dyson & Williams on gives 2×10^{-13}. My guess is the the Draine approximation is more accurate.

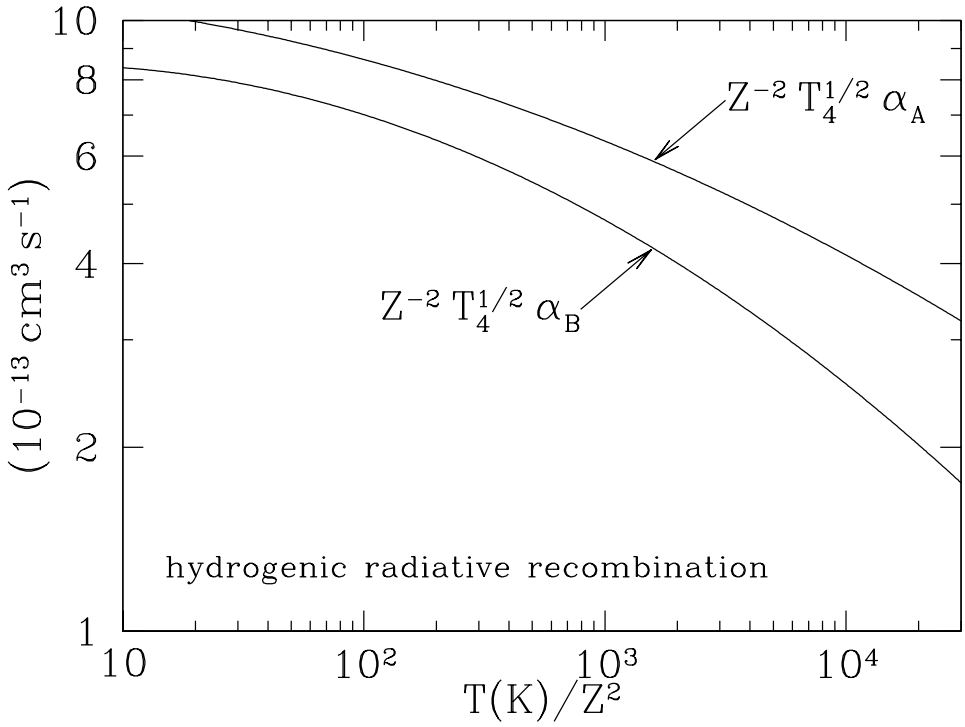
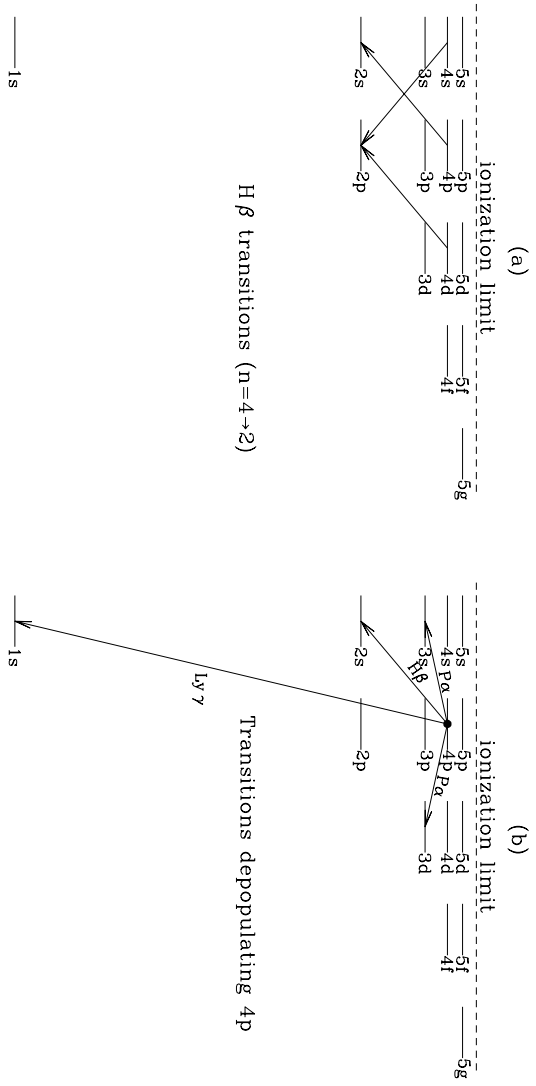


Figure 4.11: Case A and Case B recombination rate coefficients [Draine Figure 14.1].



Recombination Spectrum

Take a look at Draine Figure 14.2

We see that if radiative recombination takes place into level $4p$, it will then release energy as the electron cascades down toward the $1s$ state. The left pane shows Hβ transitions while the right shows transitions out of the $4p$ level.

These are "bound-bound" transitions due to recombination, otherwise known as recombination lines. We will discuss these further in upcoming lectures. We do have to account for some effects now, however.

In Case A recombination, it is optically thin and all radiation gets out. Easy.

In Case B recombination, it is optically thick to Lyman continuum radiation. This requires that it is also optically thick to Lyman $\alpha, \beta, \gamma...$ since the resonant absorption cross sections are much larger than the photoionization cross sections. For Case B, *all* the Lyman series is optically thick.

[What happens to the emitted photons when the radiation is optically thick?] The photons are immediately re-absorbed by nearby atoms. This is known as the **on-the-spot** assumption. We can think of emission and absorption as being from the same atom.

This means that the Lyman photons essentially are not emitted at all. Each transition to create a Lyman photon is reversed until eventually a non-Lyman photon is emitted.

This leads to a problem, however. How are we supposed to get electrons out of either the $2p$ or $2s$ states? The transition $2s \rightarrow 1s$ is strongly forbidden and only occurs with two-photon decay. This produces continuous radiation.

Since the lifetime in the $2s$ state is long, collisional de-excitation is possible. The most probable collisional de-excitation is $2s \rightarrow 2p$. In bright H II regions with high enough density, the two-photon process is suppressed by collisions.

The transition $2p \rightarrow 1s$ has a very short radiative lifetime of 1.59 ns. Collisions here are unimportant. The issue again is that for Case B recombinations, the photons will be scattered many times. Since individual ions in the WIM have velocities of $\sim 10\ \mathrm{km\,s^{-1}}$, the frequency of each emitted photon with respect to nearby atoms is not exactly at the frequency of Lyman α. The optical depth is of course lower off line center, and eventually this effect allows radiation to escape.

Radiative Recombination for Heavy Elements

Heavy element recombination is much less important than that of H and He.

For metals, we do not have to worry about recombining photons ionizing additional atoms, so we can use Case B recombination for all species.

For $T \lesssim 10^3\,\mathrm{K}$, most metals will be singly or doubly ionized, while H and He will be predominantly neutral. Most metals have low ionization potentials (Appendix D).

For $T \gtrsim 10^4\,\mathrm{K}$ most metals will be doubly ionized.

In dielectric recombination, if an ion has at least one bound electron to begin with, the combining electron can transfer energy to a bound electron, promoting the bound electron to an excited state, and removing enough energy from the first electron that it too can be captured in an excited state. The atom now has two excited electrons. This atom can "autoionize" to get rid of the energy of both excited electrons, or the electrons can relax radiatively.

Dielectric recombination is important for high temperature plasmas. This is because the

electron must have sufficient energy to produce a doubly ionized state. Some elements do have appreciable dielectric recombination rates, even in the relatively low temperature H II regions.

Ionization Balance in Collisionally Ionized Gas

We are finally now able to talk about ionization balance, when the ionization rate balances the recombination rate. Our discussion here will focus on collisionally ionized gas, i.e. the HIM. Next time we will discuss ionization balance in a radiatively ionized gas, i.e. and H II region.

When do we get collisional ionization balance? When

$$n_e \left\langle \sigma v \right\rangle_{\rm ci} n(X^{n+}) = n_e \left\langle \sigma v \right\rangle_{\rm rr} n(X^{(n+1)+}) \tag{4.44}$$

Note that this is only valid for hydrogen. For metals, we would have to include dielectric recombination too. We had the collisional ionization rate before:

$$k_{\rm ci} = \left(\frac{8kT}{\pi m_e}\right)^{1/2} \int_I^\infty \sigma_{\rm ci}(E) \frac{E}{kT} e^{-E/kT} \frac{dE}{kT} \,. \tag{4.45}$$

We can approximate the collisional cross section:

$$\sigma_{\rm ci} \approx C\pi a_0^2 \left(1 - \frac{I}{E}\right) \,, \tag{4.46}$$

where C is a constant of order unity. This expression reduces to zero when $I = E$. Draine gives a radiative recombination rate as:

$$\sigma_{\rm rr}(E) = \frac{g_\ell}{g_{\rm n}} \frac{(I+E)^2}{em_e c^2} \sigma_{\rm pi}(h\nu = I + E) \,, \tag{4.47}$$

which is called the Milne relation.

The Milne relation is derived in Draine Chapter 3. It obviously relates the ionization cross section to the radiative recombination cross section. The derivation is rather simple, and requires LTE. In LTE, the rate at which photons are removed by photoabsorption must be equal to that at which they are created by radiative recombination (withkin an energy range). Then the detailed balance,

$$\left\langle \sigma v \right\rangle_{\ell \to u} = \frac{g_u}{g_\ell} e^{(-E_{u\ell}/kT)} \left\langle \sigma v \right\rangle_{u \to \ell} \,, \tag{4.48}$$

the balance between upward and downward rate coefficients, gives us the Milne relation.

Let's introduce a new concept of oscillator strength for photoionization:

$$f_{\ell \rm u} = \frac{m_e c}{\pi e^2} \int \sigma_{\ell \rm u}(\nu) \, d\nu \,, \tag{4.49}$$

a constant multiplied by the cross section integrated over all frequencies. Oscillator strengths are strange terms to use when we have much more fundamental terms possible, but they do have one large advantage:

$$\Sigma_j f_{ij} = N\,, \tag{4.50}$$

where N is the number of electrons and the sum is over all transitions out of initial state i. The utility of the oscillator strengths is that we can easily relate the oscillator strength of a known level to that of all transitions into or out of that level.

Now, if $f_{\rm pi}$ is the oscillator strength for photoionization from the ground state, and assume $\sigma_{\rm pi} \propto (h\nu)^{-3}$ (from before), then

$$\sigma_{\rm pi}(h\nu = I) \approx \frac{2\pi e^2}{m_e c} f_{\rm pi} \frac{h}{I}\,, \tag{4.51}$$

and it can be shown that

$$\frac{\langle \sigma v \rangle_{\rm rr}}{\langle \sigma v \rangle_{\rm ci}} \approx 4\pi\alpha^3 \frac{f_{\rm pi}}{C} \frac{I}{kT} e^{I/kT}\,, \tag{4.52}$$

where α is the fine-structure constant.

Ionization balance occurs at a temperature T when the above ratio is 1:1

$$\frac{I}{kT} e^{I/kT} = \frac{C}{4\pi\alpha^3 f_{\rm pi}}\,. \tag{4.53}$$

If $C \approx 1$ and $f_{\rm pi} \approx 1/2$, $I/kT \approx 10.6$. If dielectric recombination occurs, the temperature increases.

Recombination Lines

For ionized gas, some of the bound-bound transitions we discussed previously are actually recombination lines - they result from the recombination of electrons and ions. As we saw, recombination can take place into any electron state. If the electron recombines into an excited state, it will then transition toward the ground state. The radiation it emits during this transition is called "recombination lines." This is a bit of a misnomer, since it is not the recombination per se that is causing the lines (that causes free-bound continuum of course). To detect these lines, the environment must be sufficiently so the atoms are not collisionally de-excited.

The most famous recombination line is Hα, at 656.3 nm ($n = 3 \rightarrow 2$). This line is an excellent tracer of ionized gas, and is strong enough to be used as a measure of the total ionized gas content of galaxies (via a narrow-band filter). Other famous lines are Hβ (486.1 nm, $n = 4 \rightarrow 2$), the [OIII] doublet at 459.9 nm and 500.7 nm ($^3P_2 \rightarrow^1 D_2$ and $^3P_1 \rightarrow^1 D_2$), lines of [NII], [SIII]. In images of H II regions and planetary nebulae, the red is typically Hα and the green is frequently [OIII].

Lines with $n > 40$ are in the radio regime, and are therefore "radio recombination lines" (RRLs). These lines are faint (a few Jy for the brightest H II regions, but 10 mJy more typical for faint H II regions)

As the electron cascades down the atomic levels, recombination lines are emitted at a frequency of

$$\nu_0 = R_{\rm M} c \left[\frac{1}{n^2} - \left(\frac{1}{n + \Delta n} \right) \right] , \tag{4.54}$$

where $R_{\rm M}$ is the Rydberg constant for atoms of mass M, n is the electronic level, and Δn is the change in electronic level. The Rydberg constant is not a constant at all, it is different for different elements:

$$R_{\rm M} = R_\infty \left(1 + \frac{m_{\rm e}}{M} \right)^{-1} , \tag{4.55}$$

where m_e is the electron mass, M is the nuclear mass, and

$$R_\infty = \frac{2\pi^2 m_e e^4}{ch^3} = 1.0973731568525 \times 10^5 \, {\rm cm}^{-1} . \tag{4.56}$$

Therefore, heavier elements produce the RRL transitions at lower frequencies compared to hydrogen. The frequency shift in Equation 4.54 caused by different nuclear masses (commonly called the "mass shift") is important in the study of recombination lines and plays an integral role in deriving the electron temperature of H II regions.

For $n > 40$ and $\Delta n \ll n$, Equation 4.54 becomes:

$$\nu_0 \approx \frac{2(R_{\rm M} c)\Delta n}{n^3} . \tag{4.57}$$

The spacing between adjacent lines with the same Δn is then

$$\nu_0(n) - \nu_0(n-1) \approx 2R_{\rm M} c \frac{1}{n^3 - (n-1)^3} \approx \frac{3\nu}{n} . \tag{4.58}$$

The spacing between adjacent RRLs with the same Δn increases linearly with decreasing n.

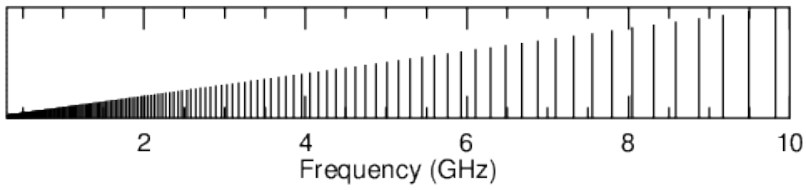


Figure 4.12: RRLs at radio frequencies. The height approximates the relative strength of the lines.

Departure Coefficients [Draine 3.8] RRLs provide a straightforward way of determining ionized gas velocities, and as we'll see they can be used to derive the temperature and density of a plasma.

In most cases we assume LTE, but that is not always a good assumption. Let's review LTE. [What can we say about a system in LTE?]

(1) Collisions dominate over radiation. This is the most basic condition of LTE, and all the rest follow from it.
(2) The velocity distribution is a Maxwellian.
(3) The ionization state is given by the Saha equation.
(4) The level populations are given by the Boltzmann factor, with $T_{\rm ex} = T$.

Deviations from LTE are caused by the radiation field. In general, we have LTE for high densities, and when the radiation field is minimized. The WIM and HIM are both at low density, so we have to take this into consideration. Ignoring this complication, we would over-estimate the line intensity and under-estimate the required temperature for the detected emission. [Gordon+Sorochenko, pg.70]

It is useful to introduce a "departure coefficient" b_n to describe the departure from LTE:

$$b_n = \frac{n[\mathrm{H}(n)]}{n_{\rm LTE}[\mathrm{H}(n)]}\,, \tag{4.59}$$

where n is the principle quantum number. When collisions dominate (in LTE), $b_n = 1$. The departure coefficient compares the true population with that expected from LTE. The true population is decreased relative to that expected in LTE due to spontaneous emission. As the density goes up, the departure coefficients tend toward unity.

What is the critical density of RRLs? Draine shows that it is

$$n_c \approx 110 \left(\frac{n_e}{10^3\ \mathrm{cm}^{-3}}\right)^{-1/7} T_4^{1/14}\,. \tag{4.60}$$

Therefore, for $n \lesssim 110$, we would expect to see departures from LTE for RRLs, for typical H II region densities.

The full definition of b_n is

$$b_n = \frac{n[\mathrm{H}(n)]}{n_{\rm LTE}[\mathrm{H}(n)]} = \frac{n[\mathrm{H}(n)]}{n_e n(\mathrm{H}^+)} \frac{(2\pi m_e kT)^{3/2}}{n^2 h^3} e^{-I_H/n^2 kT}\,. \tag{4.61}$$

We arrive at this expression from the "law of mass action." For chemical reaction $A + B \leftrightarrow C$:

$$\frac{n_{\rm LTE}(C)}{n_{\rm LTE}(A)\, n_{\rm LTE}(B)} = \frac{f(C)}{f(A) f(B)}\,, \tag{4.62}$$

where f is the partition function per unit volume. Evaluating the above equation for recombination gives the definition for the departure coefficient.

(This is just the Saha equation, but also takes into account atoms with principle quantum number n, not just the $n = 0, 1s$ state.)

Recombination Line Strengths The line absorption coefficient in LTE for RRLs is defined as

$$\kappa_\nu = \frac{c^2}{8\pi\nu_0^2} \frac{g_{\rm u}}{g_\ell} n_\ell A_{{\rm u}\ell} \left(1 - \frac{n_{\rm u}/g_{\rm u}}{n_\ell/g_\ell}\right) \phi(\nu)\,, \tag{4.63}$$

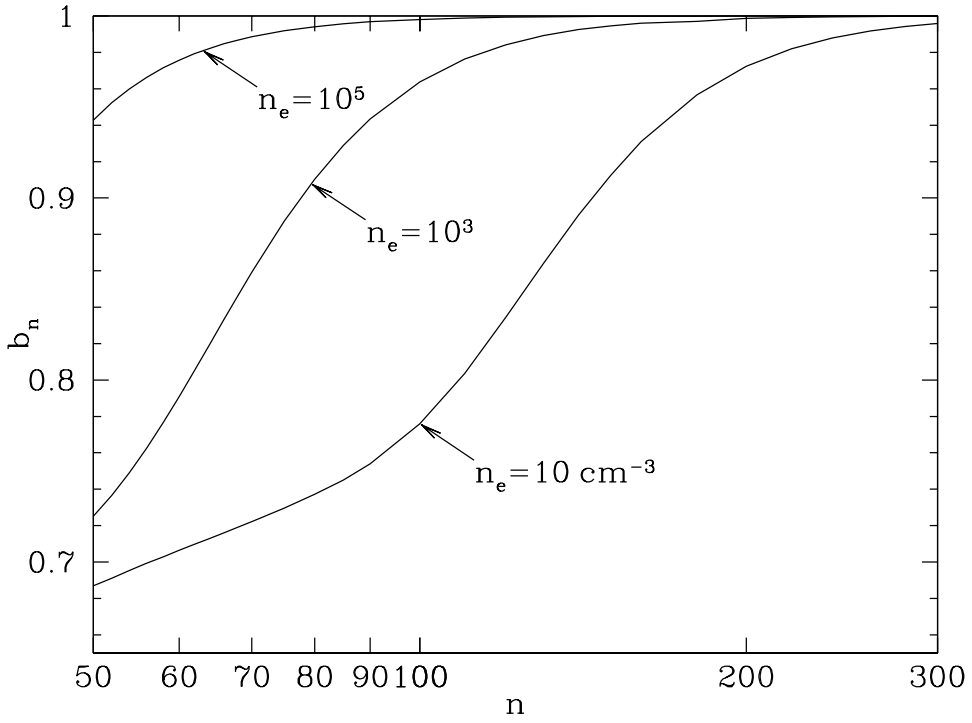


Figure 4.13: Departure coefficients.

where $g_{\rm u}$ and g_ℓ are the statistical weights for levels u and ℓ, n_ℓ is the number density in level ℓ, $A_{{\rm u}\ell}$ is the Einstein A coefficient of spontaneous emission from level u to level ℓ, and $\phi(\nu)$ is the normalized line profile. The degeneracies for hydrogen (and all hydrogenic atoms) are $2n^2$.

The negative term in the parentheses accounts for stimulated emission. This term is absent in LTE calculations of course.

From our definition of the departure coefficient, $n_n/g_n \propto b_n \exp(I_H/n^2kT)$. For α transitions therefore,

$$\left(1 - \frac{n_{\rm u}/g_{\rm u}}{n_\ell/g_\ell}\right) = 1 - \frac{b_{n+1}}{b_n} \exp\left[-\frac{(2n+1)}{n^2(n+1)}\frac{I_H}{kT}\right] \tag{4.64}$$

Salem & Brocklehurst define a new parameter β_n:

$$\beta_n = \frac{1 - \frac{n_{\rm u}/g_{\rm u}}{n_\ell/g_\ell}}{1 - \exp(-h\nu/kT)} \tag{4.65}$$

Therefore,

$$\kappa_\nu = \frac{c^2}{8\pi\nu_{n+1,n}^2} n_n A_{n+1,n} \phi(\nu) \beta_n \left(1 - e^{-h\nu/kT}\right) . \tag{4.66}$$

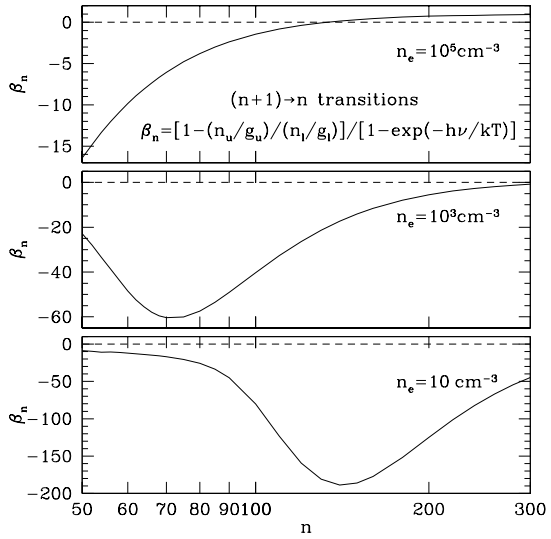


Figure 4.14: Draine Figure 10.3

RRLs are always optically thin, so we know that $T \propto \tau \propto \kappa$. When dealing with LTE The non-LTE line absorption coefficient is usually denoted κ, while κ^* is actually the LTE term.

If β_n is not unity, we have stimulated emission, and in extreme cases RRLs can be masing.

The total RRL line strengths for α lines is $I \propto \tau \propto \kappa \propto \nu^{-1} A_{n+1,n}$. For RRLs, $A \propto n^{-5}$ (see Essential radio astronomy course), but in the radio regime this factor makes little difference. We can therefore say that for lines of a given Δn, the line strength is inversely proportional to the frequency.

It is important to remember that the intensity is misleading. We know that Hα is the strongest recombination line, but its frequency is quite high! This is because the flux density $S \propto \nu^2 T$, so $S \propto \nu$. Flux density is the real quantity of interest.

4.5 Propagation of EM Radiation through a Plasma

Light in a vacuum travels at the speed of light. The ISM is not a vacuum, and the speed light travels is reduced. This effect is small and dependent on the number density of free electrons. Because the path lengths are enormous, however, the small effects add up to a measurable effect: light is delayed as it passes through the ISM and this delay is frequency dependent.

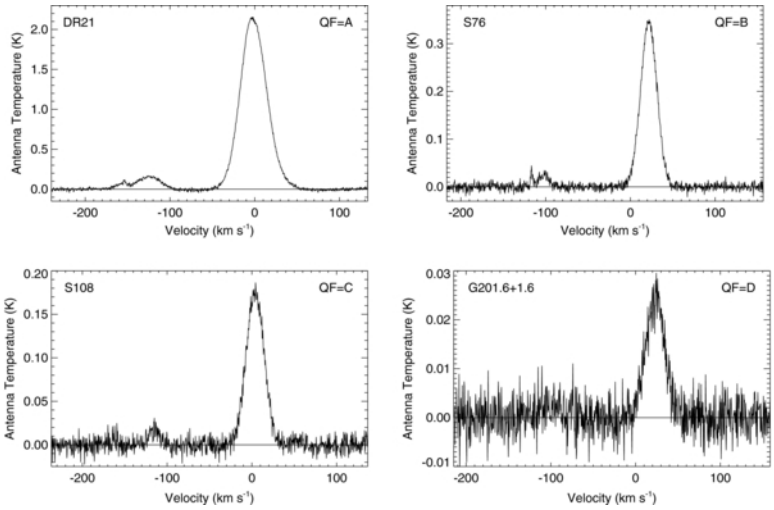


Figure 4.15: From Balser et al., 2011.

Just for a frame of reference, Draine notes that for a pulsar 3 kpc away, assume the dispersion measure $DM \approx 10^2$ cm^{-3} pc (see below). At 1 GHz, the time delay will be ~ 0.4 s after traveling at the speed of light for 10^4 yr. Draine lists the mean value for n_e in the ISM is ~ 0.05 cm^{-3}, although I have seen 0.03 cm^{-3} used also.

The following notes are from Michael Lam when he guest lectured for me in 2016.

We consider the oscillation of cold (the electron thermal speeds are negligible), free electrons in a propagating electric field. For a single electron, we have

$$m_e \ddot{\mathbf{r}} = e\mathbf{E}. \tag{4.67}$$

The current density is

$$\mathbf{J} = \rho \mathbf{v} = n_e e \dot{\mathbf{r}} \tag{4.68}$$

and its time derivative is

$$\frac{\partial \mathbf{J}}{\partial t} = \frac{\partial}{\partial t}\left(n_e e \dot{\mathbf{r}}\right) = \left(n_e e \ddot{\mathbf{r}}\right) = n_e e \left(\frac{e\mathbf{E}}{m_e}\right) = \frac{n_e e^2}{m_e}\mathbf{E}. \tag{4.69}$$

Let's now assume we have a linearly-polarized plane wave propagating in the x-direction. Therefore, we can write the electric field as

$$\mathbf{E} = E_0 \exp\left[i(kz - \omega t)\right] \hat{\mathbf{x}}. \tag{4.70}$$

This equation satisfies the wave equation, so it will be useful later to determine the two quantities:

$$\nabla^2 \mathbf{E} = \frac{\partial^2}{\partial z^2} E_0 \exp\left[i(kz - \omega t)\right] \hat{\mathbf{x}} = -k^2 E_0 \exp\left[i(kz - \omega t)\right] \hat{\mathbf{x}} = -k^2 \mathbf{E} \tag{4.71}$$

$$\frac{\partial^2 \mathbf{E}}{\partial t^2} = \frac{\partial^2}{\partial t^2} E_0 \exp\left[i(kz - \omega t)\right] \hat{\mathbf{x}} = -\omega^2 E_0 \exp\left[i(kz - \omega t)\right] \hat{\mathbf{x}} = -\omega^2 \mathbf{E} \tag{4.72}$$

We will also use three of Maxwell's equations in Gaussian (cgs) units:

$$\begin{aligned}
(\text{Gauss})\nabla\cdot\mathbf{E} &= 4\pi\rho && (4.73)\\
(\text{Faraday})\nabla\times\mathbf{E} &= -\frac{1}{c}\frac{\partial\mathbf{B}}{\partial t} && (4.74)\\
(\text{Ampere})\nabla\times\mathbf{B} &= \frac{1}{c}\left(4\pi\mathbf{J}+\frac{\partial\mathbf{E}}{\partial t}\right). && (4.75)
\end{aligned}$$

We start by taking the curl of both sides of Eq. 4.74 and then pluging in values appropriately

$$\begin{aligned}
\nabla\times(\nabla\times\mathbf{E}) &= \nabla\times\left(-\frac{1}{c}\frac{\partial\mathbf{B}}{\partial t}\right) && (4.76)\\
\rightarrow \nabla(\nabla\cdot\mathbf{E})-\nabla^2\mathbf{E} &= -\frac{1}{c}\frac{\partial}{\partial t}(\nabla\times\mathbf{B}) && (4.77)\\
\rightarrow \nabla(4\pi\rho)-\left(-k^2\mathbf{E}\right) &= -\frac{1}{c}\frac{\partial}{\partial t}\left[\frac{1}{c}\left(4\pi\mathbf{J}+\frac{\partial\mathbf{E}}{\partial t}\right)\right] && (4.78)\\
\rightarrow 4\pi\cancel{\nabla\rho}+k^2\mathbf{E} &= -\frac{4\pi}{c^2}\frac{\partial\mathbf{J}}{\partial t}-\frac{1}{c^2}\frac{\partial^2\mathbf{E}}{\partial t^2} && (4.79)\\
\rightarrow k^2\mathbf{E} &= -\frac{4\pi}{c^2}\frac{n_e e^2}{m_e}\mathbf{E}+\frac{\omega^2}{c^2}\mathbf{E} && (4.80)\\
\implies k^2c^2 &= -\frac{4\pi n_e e^2}{m_e}+\omega^2 && (4.81)
\end{aligned}$$

We can define the (angular) plasma frequency as

$$\omega_p^2 \equiv \frac{4\pi n_e e^2}{m_e} \tag{4.82}$$

and therefore we arrive at the dispersion relation

$$\omega^2 = k^2c^2+\omega_p^2. \tag{4.83}$$

The plasma frequency is related to the electron density as

$$\nu_p = \frac{\omega_p}{2\pi} \approx 8.979 \text{ kHz } \left(\frac{n_e}{\text{cm}^{-3}}\right)^{1/2}. \tag{4.84}$$

The propagation speed is given by the group velocity

$$
\begin{aligned}
v_g \equiv \frac{\partial \omega}{\partial k} &= \frac{\partial}{\partial k}\omega && (4.85)\\
&= \frac{\partial}{\partial k}\left(k^2c^2+\omega_p^2\right)^{1/2} && (4.86)\\
&= \frac{\cancel{2}kc^2}{\cancel{2}\left(k^2c^2+\omega_p^2\right)^{1/2}} && (4.87)\\
&= \frac{\cancel{c}\left(\omega^2-\omega_p^2\right)^{1/2}\cancel{c}}{\omega} && (4.88)\\
&= c\left(1-\frac{\omega_p^2}{\omega^2}\right)^{1/2} && (4.89)\\
&= c\left(1-\frac{\nu_p^2}{\nu^2}\right)^{1/2} && (4.90)\\
&\equiv c\mu && (4.91)
\end{aligned}
$$

where $\mu \leq 1$ is the index of refraction. Below the plasma frequency, μ is imaginary and the waves cannot propagate. For the ionosphere, the electron density peaks at about 10^6 cm^{-3} and so the plasma frequency is about 9 MHz. In the ISM, for $n_e \sim 0.1$ cm^{-3}, the plasma frequency is about 3 kHz.

4.5.1 Dispersive Time Delay

The total propagation time as a function of path length through the medium is

$$
\begin{aligned}
t_{\rm total} &= \int_0^D \frac{dl}{v_g} && (4.92)\\
&= \int_0^D \frac{dl}{c}\left(1-\frac{\nu_p^2}{\nu^2}\right)^{-1/2} && (4.93)\\
&\approx \int_0^D \frac{dl}{c}\left(1+\frac{\nu_p^2}{2\nu^2}\right) && (4.94)\\
&= \int_0^D \frac{dl}{c} + \int_0^D \frac{dl}{c}\frac{\nu_p^2}{2\nu^2} && (4.95)\\
&= \frac{D}{c} + \frac{e^2}{2\pi m_e c}\frac{\int_0^D n_e(l)dl}{\nu^2} && (4.96)\\
&= t_{\rm geometric} + t_{\rm dispersive}. && (4.97)
\end{aligned}
$$

where in the last step we broke up the total time into the geometric travel time and the dispersive delay. Therefore,

$$t_{\rm dispersive} = \frac{e^2}{2\pi m_e c}\frac{\int_0^D n_e(l)dl}{\nu^2} \quad (4.98)$$

$$\equiv K\frac{\rm DM}{\nu^2} \quad (4.99)$$

$$\approx 4.149 \text{ ms } \left(\frac{\rm DM}{\rm pc\ cm^{-3}}\right)\left(\frac{\nu}{\rm GHz}\right)^{-2} \quad (4.100)$$

where $\mathrm{DM} \equiv \int_0^D n_e(l)dl$ is the dispersion measure and $K \equiv \frac{e^2}{2\pi m_e c} \approx 4.149\,\mathrm{ms\ GHz^2\ pc^{-1}\ cm^3}$ is the dispersion constant.

4.5.2 Scintillation

The intensity of light we receive from astronomical objects is often found to vary over short time scales. This is called **scintillation**, and is the technical term for the twinkling of stars.

Scintillation results from small-scale changes in the medium between the observer and the source. The twinkling of stars is caused by small-scale changes in our atmosphere due to turbulence. We also see scintillation in the intensity of pulsars and background radio sources. Scintillation is caused by turbulence in the medium.

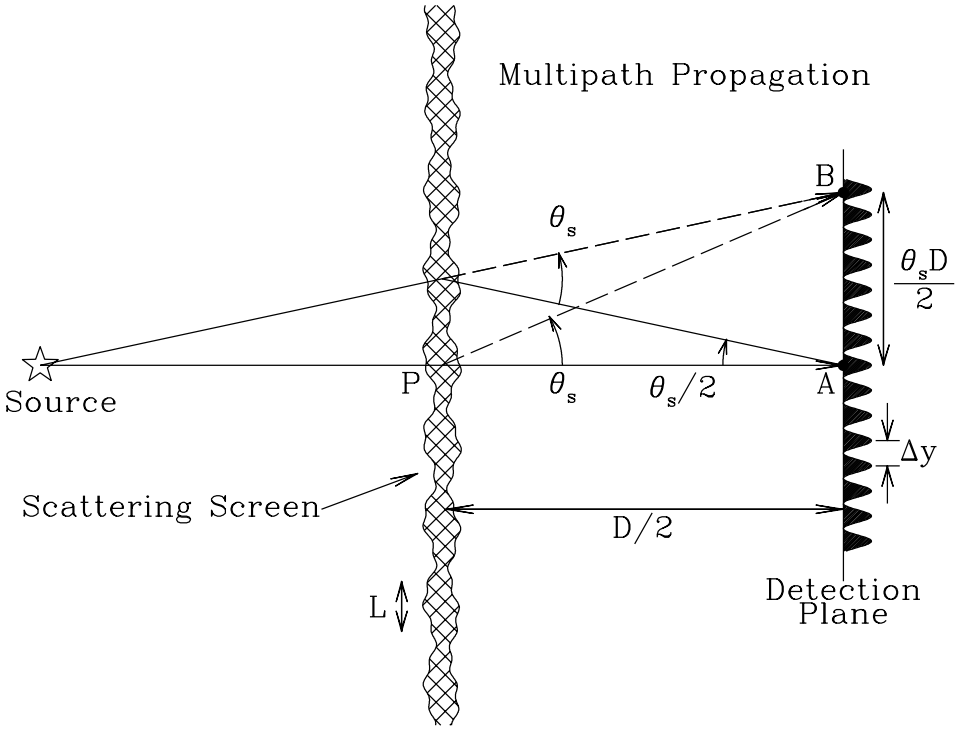
Consider EM radiation scattering off a "screen" of electrons. If the scattering angle is sufficiently large, there are multiple possible paths for the radiation to take to get to the observer. The radiation from these multiple paths can constructively and destructively interfere, leading the intensity variations.

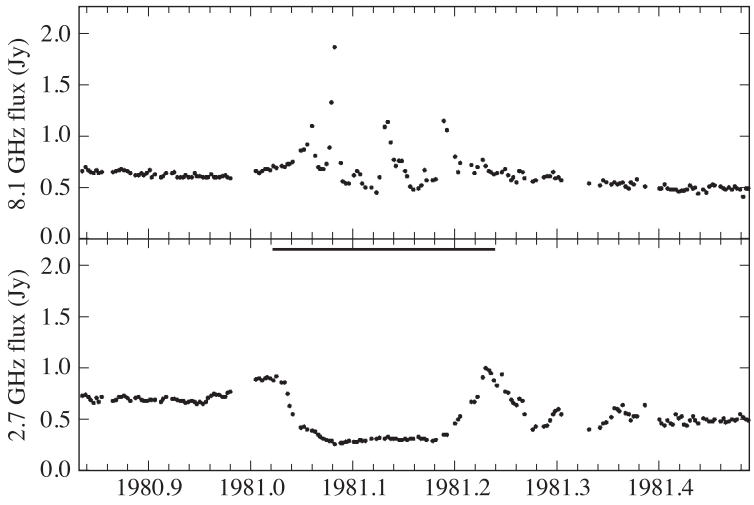
Scintillation is minimized for sources at small distance and at high frequency.

"Extreme" scattering events have also been observed toward extragalactic radio sources. The rotation of the Galaxy occasionally causes material to pass between background radio sources and ourselves. Sometimes there is a little plasma blob in between us and the source. Scattering off this plasma blob results in extreme scattering events where the brightness increases by a factor of ~ 2.

The plasma blobs must be small to agree with the timescale of intensity enhancements, but we still don't have a good model for how they can operate. Analysis of the likely blob properties indicates that they are not in pressure equilibrium with their surroundings, and should actually expand. If this is correct, they must be very transient, which would explain their rarity.

[From Cal. Tech course http://www.its.caltech.edu/∼kamion/Ay126/Bfields.pdf]


Multipath Propagation
B
θ_s
$\frac{\theta_s D}{2}$
Source
P
θ_s
$\theta_s/2$
A
Δy
Scattering Screen
D/2
L
Detection Plane


8.1 GHz flux (Jy)
2.7 GHz flux (Jy)
0.0
0.5
1.0
1.5
2.0
1980.9
1981.0
1981.1
1981.2
1981.3
1981.4

4.5.3 Faraday rotation

If the plasma is magnetized, then there is an additional effect that acts on the polarization of the electromagnetic wave. The Faraday rotation effect appears during the propagation of electromagnetic waves in a magnetized plasma.

A linearly polarized wave can be decomposed into opposite-handed circularly polarized components. The right-handed and left-handed circularly polarized waves propagate with different phase velocities within the magneto-ionic material. Draine says

$$k^2c^2 = \omega^2 - \frac{\omega_p^2}{1 \pm \omega_B/\omega}\,, \tag{4.101}$$

where the cyclotron frequency is

$$\omega_B \equiv \frac{eB_\parallel}{m_e c} \tag{4.102}$$

and the $\pm$ refers to the two circular polarizations.

This effectively rotates the plane of polarization of the electromagnetic wave. Suppose there is a magnetic field in the plasma directed along the line of sight. Electrons will then spiral in one particular direction around this magnetic field, and so the indexes of refraction for right- and left-circularly polarized electromagnetic waves will differ. The propagation speeds for right- and left-circularly polarized waves will therefore differ slightly. Recalling that a linearly-polarized wave is a superposition of two circularly-polarized waves, the linearly-polarized wave will undergo something like a beat phenomenon that occurs when two waves of slightly different frequencies are superposed. What this results in is a rotation of the linear polarization of a linearly polarized EM wave by an angle, $\Psi = RM\lambda^2$, where λ is the wavelength and

$$RM = \frac{1}{2\pi}\frac{e^3}{m_e^2c^4}\int_0^L n_e B_\parallel dL \tag{4.103}$$

If the DM and RM are both measured, after evaluating the constants the electron-density-weighted mean line-of-sight magnetic field is

$$\langle B_\parallel \rangle = \frac{RM}{8.12\times10^{-5}\mathrm{radcm}^{-2}}\frac{\mathrm{cm}^{-3}\ \mathrm{pc}}{DM}\,\mu\mathrm{G}\,. \tag{4.104}$$

This can be measured along many different lines of sight, and also to pulsars at different distances along similar lines of sight, to get information about the three-dimensional magnetic field. Measurements indicate magnetic fields $B \sim 2-4\,\mu$G in the spiral arms and slightly smaller in the interarm regions, with a sign flip between arm and interarm.

Faraday depolarization removes the relationship between the rotation measure and λ^2. It is

$$DP = \left(\frac{PI_1}{PI_2}\right)\left(\frac{\nu_2}{\nu_1}\right)^\alpha\,, \tag{4.105}$$

where α is the synchrotron spectral index.

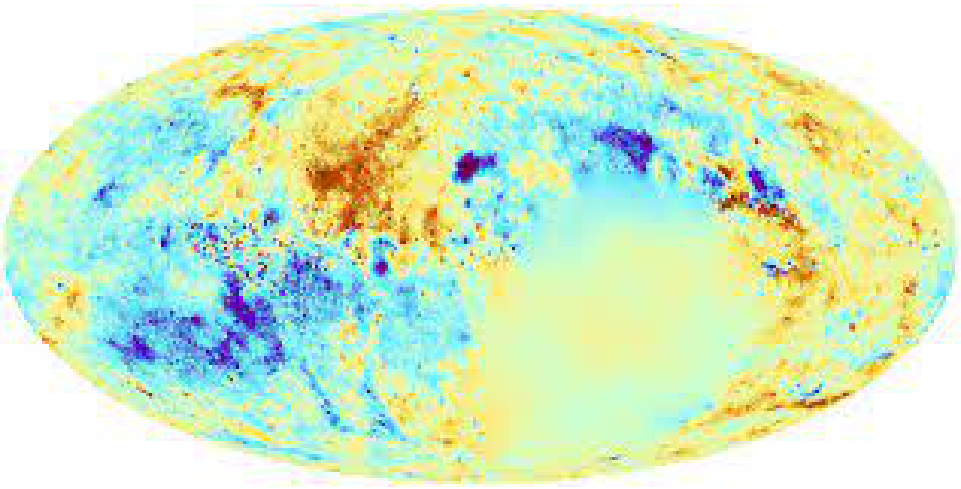


Figure 4.16: Faraday rotation measured over entire sky (from Hutschenreuter et al., 2021).

4.6 HII Regions

There is unanimous agreement in the astronomical community that H II regions are the most interesting and important objects in the Universe. Let's spend a little more time understanding their physics.

H II regions are technically any ionized region of the ISM. In practice, however, this term is reserved for the plasma spheres surrounding early-type (OB) stars. OB stars emit much of their energy in the UV portion of the spectrum, and thus can ionize their surrounding medium. The cutoff point between which stars can ionize their surroundings appreciably, and which cannot is about B2. This is somewhat arbitrary though as even the Sun can ionize its surroundings.

When talking about H II regions we often think only about main sequence stars. Evolved stars can also create H II regions, although these are more rare due to the shorter lifetimes in these evolved states.

4.6.1 Strömgren Spheres

Strömgren (1939) derived the size of an idealized, fully ionized, pure hydrogen, spherical plasma zone surrounding an H II region. This turns out to be not such a bad set of assumptions!

In an H II region, we should have ionization balance: the number of ionizations per second should balance the number of recombinations per second. If this were not the case, nebulae would be growing or shrinking. The number of hydrogen ionizing, or "Lyman continuum", photons is Q_0 (also $N_{\rm ly}$ sometimes used). Equating ionization and recombination:

$$Q_0 = \frac{4\pi}{3} R_S^3 \alpha_B n(H^+) n_e \simeq \frac{4\pi}{3} R_S^3 \alpha_B n_e^2. \tag{4.106}$$

Therefore,

$$R_s = \left(\frac{3Q_0}{4\pi n_e^2 \alpha_B}\right)^{1/3} = 3.17 Q_{0,49}^{1/3} n_2^{-2/3} T_4^{0.28}\,\mathrm{pc}\,. \tag{4.107}$$

From theoretical calculations, $Q_{0,49}$ corresponds to the emission from an O6V star (which for reference is the spectral typr of the largest star in Orion's Trapezium). We see that Strömgren spheres should be on the order of pc. This matches observations (e.g., again, of Orion).

What assumptions go into this derivation?
(0) Ionization exactly balances recombination.
(1) The nebula is entirely hydrogen.
(2) The ionized fraction in the nebula is 100%.
(3) The boundary between the nebula and the surrounding medium is sharp (otherwise, what would be the meaning of a Stromgren sphere?).
(4) The nebula can be characterized by a single density.
We must assume the first assumption is valid, but we can examine some of the others.

4.6.2 Helium Ionization

The above assumes that the nebula is pure hydrogen. What effect would helium have on the size of an ionized region?

The recombination rate for He is ~ 1.9 times greater than that of H (Draine chapter 14, although I wasn't able to work out how exactly he got this number, it's likely just from the recombination rates), for $T \simeq 10^4$ K. Since the abundance of He is $\sim 10\%$ that of H by number, the volume rate for He recombinations will be $\sim 18\%$ that of H.

The first ionization potential for He is 24.6 eV. We therefore need a relatively hard radiation field to ionize He (thus the assumption of a pure H nebula above). Smaller stars that are able to create HII regions nevertheless may not be able to ionize He appreciably. Q_1 is the rate for He ionizing photons. If $Q_1 < 0.15 Q_0$, the He ionization zone will be smaller than the H ionized zone. If $Q_1 > 0.15 Q_0$, they will have the same volume, and He will be singly ionized and H will be fully ionized. This corresponds to an O6V star (Draine table 15., Figure 15.5). In no case will the He ionized zone be larger than that of H, since no stars have $Q_1 > Q_0$.

If the ionized spheres completely overlap, the ratio of the number of helium ions must equal the ratio of the helium and hydrogen densities:

$$N(\mathrm{He}^+)/N(\mathrm{H}^+) = n_\mathrm{H}/n_\mathrm{He} \approx 0.1\,. \tag{4.108}$$

Since $n = N/Volume$, $R \propto (N/n)^{1/3}$ and

$$\frac{R_\mathrm{He}}{R_\mathrm{H}} = \left(\frac{N_\mathrm{He}}{N_\mathrm{H}}\frac{n_\mathrm{H}}{n_\mathrm{He}}\right)^{1/3}\,, \tag{4.109}$$

if $N(\mathrm{He}^+)/N(\mathrm{H}^+) < n_\mathrm{He}/n_\mathrm{H}$ the helium ionized zone will be smaller than that of hydrogen.

Draine derives the relationship between these ratios and the hardness of the radiation field.

$$\frac{N(\mathrm{He}^+)}{N(\mathrm{H}^+)} \approx \frac{0.68(Q_1/Q_0)}{1-0.17(Q_1/Q_0)}\,. \tag{4.110}$$

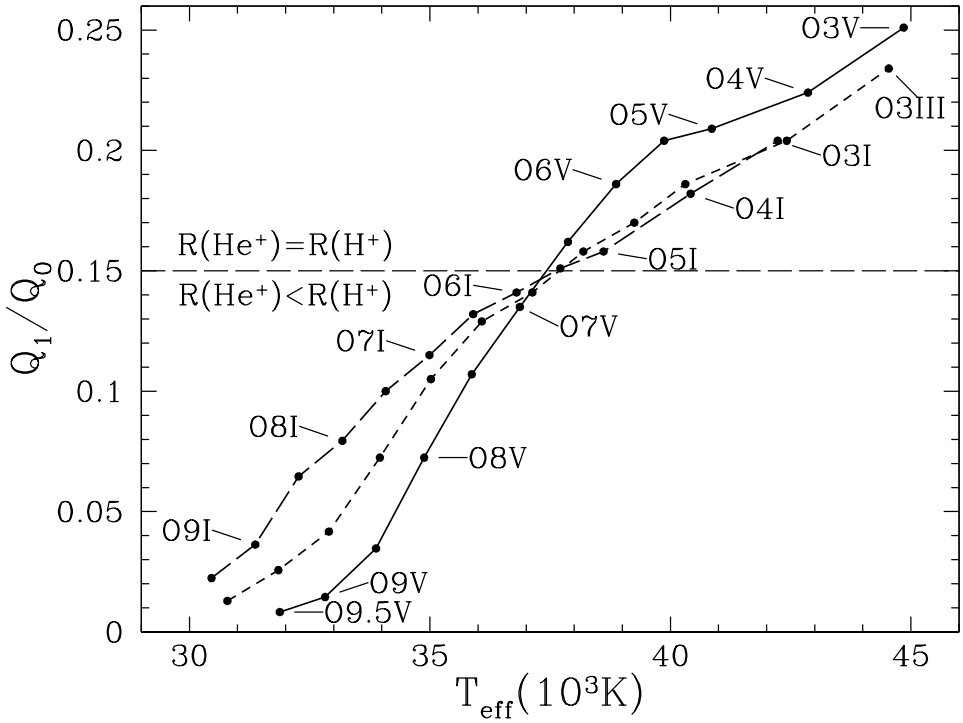


Figure 4.17: Draine 15.5. Ratios of the hydrogen and helium ionization rates, and zone sizes.

If helium is ionized, of course it is removing photons from the nebula that would otherwise ionize hydrogen. This decreases the size of the hydrogen ionized region.

4.6.3 Fractional Ionization [see Dyson & Williams]

We assumed that the neutral fraction in a Stromgren sphere is exactly zero. Otherwise, the region would be larger since ionizing photons could travel further. Is this a good assumption? How ionized are H II regions? Is there a neutral fraction within them?

In the steady state condition, again assume ionization balance. In this case:

$$n_e^2 \alpha_B x^2 = \frac{Q(r)}{4\pi r^2} n_e (1-x)\sigma_{\mathrm{pi}}\,, \tag{4.111}$$

where x is the ionization fraction ($x = 1$ is fully ionized), r is the distance from the center of the nebula, and σ_{pi} is the photoionization cross section.

We are assuming here that the radiation just falls of as r^{-2}.

Rearranging, and inserting $4/3\pi r^3\alpha_B = Q(0)$,

$$\frac{x^2}{1-x} = \frac{Q(r)}{Q(0)}\frac{n_e\sigma_{\rm pi}R_s}{3y^2}\,, \tag{4.112}$$

where $y = r/R_s$.

Putting in reasonable values here, we find that $x \simeq 1$ and $x^2 \approx 1$. Therefore, it is a good assumption that H II regions are completely ionized.

4.6.4 The Boundary of a Nebula [D+W 5.2.7]

Our derivation of R_s assumed that $x \approx 1$ everywhere in the nebula. This obviously cannot be the case *everywhere*, as there must be a transition zone with the neutral medium outside the H II region. How sharp-edged is this transition? If it is not very sharp, what is the physical meaning of a Strömgren radius?

Let's define the flux of ionizing photons $J = Q(0)/4\pi r^2$. We know that the flux at $r + dr$ is $J + dJ$, and so

$$J + dJ = J - n_H\sigma_{\rm pi}Jdr\,. \tag{4.113}$$

Since $n_H = n(1-x)$,

$$\frac{dJ}{dr} = n(1-x)\sigma_{\rm pi}J\,. \tag{4.114}$$

We want to eliminate J from the equation, which we can do easily with our previous relation for ionization balance:

$$\frac{x^2}{x-1} = \frac{J}{n}\frac{\sigma_{\rm pi}}{\alpha_B}\,, \tag{4.115}$$

which gives

$$\frac{2x(x-1)-x^2}{(x-1)^2}\frac{dx}{dr} = \frac{2-x}{x(1-x)^2}\frac{dx}{dr} = \frac{dJ}{dr}\frac{\sigma_{\rm pi}}{\alpha_B}\,, \tag{4.116}$$

$$\frac{dx}{dr} = \sigma_{\rm pi}n\frac{x(1-x)^2}{2-x}\,. \tag{4.117}$$

This equation can be easily solved, and we find that the degree of ionization falls really quickly, over a distance of $\sim 10(\sigma_{\rm pi}n)^{-1}$, or a few mfps. This is a very small distance, so we can say that ideal H II regions are indeed sharp-edged.

One caveat to this derivation is that real H II regions don't expand into a homogeneous medium. Therefore, some directions will have a larger extent than others. Although sharp-edged in each direction, real H II regions are not necessarily spherical.

4.6.5 Evolution of an H II Region

The Strömgren radius is not reached instantly. H II regions expand as they age. This expansion is typically characterized by two phases: (1) expansion of the H II region's size to

reach the Strömgren radius, and (2) expansion beyond the Strömgren radius until the H II region attains pressure equilibrium with the ambient gas.

The first phase of H II region evolution begins when a young, massive OB star first begins to produce UV photons. The ionization of neutral hydrogen surrounding the young star creates a region of ionized gas. The boundary between this ionized gas and the neutral medium, the ionization front, expands rapidly to reach the Strömgren radius as the surrounding hydrogen is ionized. The characteristic time for the expansion to reach the Strömgren radius is the recombination timescale (From Spitzer):

$$t = (n_n \alpha_B)^{-1} \,, \tag{4.118}$$

where n_n is the ambient density of the neutral medium. For a characteristic density at this phase, ($n_n \approx 10^5$), this initial expansion would take only a few years. [This doesn't seem to jibe with observations, but is maybe only illustrative...] The expansion of the Strömgren radius is supersonic. Therefore the density inside and outside the ionization front are unable to equilibrate. Until the H II region reaches the size of the Strömgren radius, the ionized and neutral gas are essentially at rest with respect to one another, although the ionization front itself is of course expanding.

The second phase of H II region evolution begins when the H II region's size reaches approximately the Strömgren radius. During photo-ionization, the number of particles within the H II region doubles. Coupled with the increased temperature inside H II regions, this creates a pressure inbalance and leads to the expansion of the H II region ionization front (Oort, 1954). At this stage the rate of expansion of the ionization front drops to the sound speed of the ionized gas and a shock front is created. This shock front moves supersonically into the ambient medium, ahead of the ionization front. The time-evolution of the ionization front radius is given by Spizer (1978):

$$R_i = R_s \left(1 + \frac{7Ct}{4R_S}\right)^{4/7} \,. \tag{4.119}$$

where R_i is the H II region radius, R_s is the Strömgren radius, C is the sound speed in the ionized material, and t is the elapsed time since the Strömgren radius formed in seconds. An H II region will cease its expansion when the pressure of the ionized gas is equal to that of the ambient medium:

$$2n_i k T_i = n_n k T_n \,, \tag{4.120}$$

where n_i, T_i and n_n, T_n are the is the gas density and temperature of the ionized and ambient mediums, respectively, In pressure equilibrium, the final H II region radius, R_f, is:

$$R_f = R_s(2T_i/T_n) \,. \tag{4.121}$$

Under the assumption of ionization balance, and for typical H II region values, the final radius of an H II region is about 30 times the Strömgren radius, $R_f \simeq 30R_s$.

It is very difficult for H II regions to achieve pressure equilibrium with the ambient interstellar medium (ISM) because OB stars are too short-lived. Therefore, H II regions expand their

entire lifetimes. Smaller H II regions should statistically be younger and larger H II regions should statistically be the oldest. Note that this treatment is simplistic! Self-gravity will also impeded expansion, but was not accounted for.

http://adsabs.harvard.edu/abs/2011ApJ...733...16I
http://adsabs.harvard.edu/abs/1989ApJS...69..831W

4.6.6 Nebular Diagnostics

[What do level populations depend on? Density and temperature, which combine to set the collision rate/energy per collision!] This can be really powerful. If we can determine the level populations through measurement of the line strengths, we can back out the density and temperature.

To be useful, we need abundant elements with accessible energy levels that can be collisionally excited, and from which we can observe emission or absorption lines. Luckily, there are plenty of such elements. This discussion will focus on plasma, since at the temperature and density of neutral gases these conditions are not met.

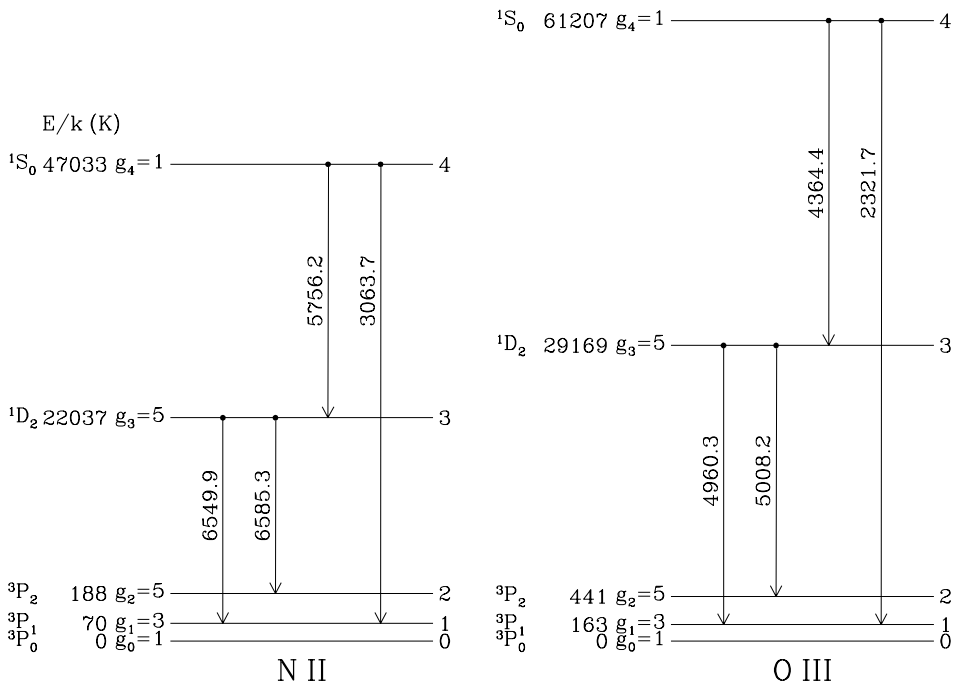
There are two main things we want to learn: temperature and density. The best lines to use for temperature determinations come from elements that have large differences in energy between excited states comparable to kT. [Why?] The energy (temperature) and frequency (density) of the collisions will be important. The best lines to use for density determinations are those with closely-spaced energy levels. Each collision, regardless of energy, will populate a level. The temperature of the plasma will not be a large factor, only the frequency of collisions (density).

4.6.7 Temperature Diagnostics Using Optical/UV Lines

Six-electron atoms and ions turn out to be very important for temperature determinations of 10^4 K plasma. $2p^2$ is the lowest configuration. The lowest term is ^{3}P, the first excited terms are ^{1}D and ^{1}S. If the ^{1}S term is at low enough energy ($E/k \lesssim 70,000$ K) so that collisional ionization in 10^4 K gas is not prohibitively slow, and the abundance is not too low, we get emission from the ^{1}D and ^{1}S levels. The relative strengths of the lines are very sensitive to temperature, because the difference in energy levels is great. Therefore, at low temperatures, collisions will not efficiently populate ^{1}S, no matter how frequent. At high temperatures, they will.

What are 6-electron atoms and ions? CI, NII, OIII, FIV, NeV, etc. C is easily ionized and so is relatively rare. FIII and NeIV have high ionization potentials so it's hard to get to FIV and NeV. So NII and OIII it is! Similar arguments apply to 8/14/16-electron atoms/ions.

Consider a three-level 6-electron ion. In the low density limit, all ions will be in the ground state $^3P_0$. Every collisional excitation will be followed by radiative de-excitation, at low densities (high densities will collisionally de-excite). The "branching ratios" are determined by the Einstein A coefficients. For example, the probability of a $4 \rightarrow 3$ transition is $A_{43}/(A_{41} + A_{43})$



(the $4 \to 2$ and $4 \to 0$ transitions are forbidden by $\Delta J = 2$ and $\Delta J = 0$, so their Einstein As are tiny and can be ignored). The power radiated per unit volume for the $4 \to 3$ and $3 \to 2$ transitions is:

$$P(4 \to 3) = E_{43}[n_0 C_{04}] \frac{A_{43}}{A_{43} + A_{41}} \tag{4.122}$$

$$P(3 \to 2) = E_{32} \left[n_0 C_{03} + n_0 C_{04} \frac{A_{43}}{A_{43} + A_{41}} \right] \frac{A_{32}}{A_{32} + A_{31}} \,, \tag{4.123}$$

where the C terms are collisional excitation rates in units of $\mathrm{cm}^{-3}\ \mathrm{s}^{-1}$. Note that for a higher number of levels, these expressions will get complicated! Therefore, $n_0 C$ is a rate in s^{-1}. The $n_0 C_{03}$ term gives collisions directly to the 3 level, while the $n_0 C_{04} \frac{A_{43}}{A_{43}+A_{41}}$ term gives radiative transitions into the 3 level. The ion collision rates are given by the expression (Draine Chapter 2):

$$C_{\ell \mathrm{u}} = 8.629 \times 10^{-8} T_4^{-1/2} \frac{\Omega_{\ell \mathrm{u}}}{g_\ell} e^{-E_{\mathrm{u}\ell}/kT} n_e \ \mathrm{cm}^{-3}\ \mathrm{s}^{-1} \,, \tag{4.124}$$

where $\Omega_{\ell \mathrm{u}}$ is the "collision strength," a QM-calculated term. We are dealing here with electron-ion collisions. [Why? Electrons are energetic and fast-moving.]

We can therefore calculate theoretical line ratios for a range of densities, and compare those to that observed.

Draine mentions that in the limit that $n_e \to 0$, the emissivity ratio is

$$\frac{j(4 \to 3)}{j(3 \to 2)} = \frac{A_{43} E_{43}}{A_{32} E_{32}} \frac{(A_{32} + A_{31}) \Omega_{04} e^{-E_{43}/kT}}{[(A_{43} + A_{41}) \Omega_{03} + A_{43} \Omega_{04} e^{-E_{43}/kT}]} \,. \tag{4.125}$$

Here we are using emisivities. The emissivity is $j_\nu = 1/(4\pi) n_{\rm u} A_{{\rm u}\ell} h\nu \phi_\nu$, and has units of power per unit frequency per unit solid angle per unit volume. How did Draine derive the above equations? We want to know the density $n_{\rm u} = n_\ell C_{{\rm u}\ell}$. In LTE, we know that the source function $S_\nu = j_\nu/\kappa_\nu = B_\nu(T_{\rm ex})$. We are free to use j or κ to understand fluxes and intensities, but we'll use j here.

OK, so the above is a little complicated as far as expressions go. The point here is that only atomic physics go into the ratio. If we know the atomic physics, a measurement of the line ratio will give you the temperature.

This only applies in the low density limit. How low density? Below the critical density is a good guess! So each line ratio will have a range of densities over which it is appropriate. At higher densities, the line ratio depends on the density itself, which complicates matters.

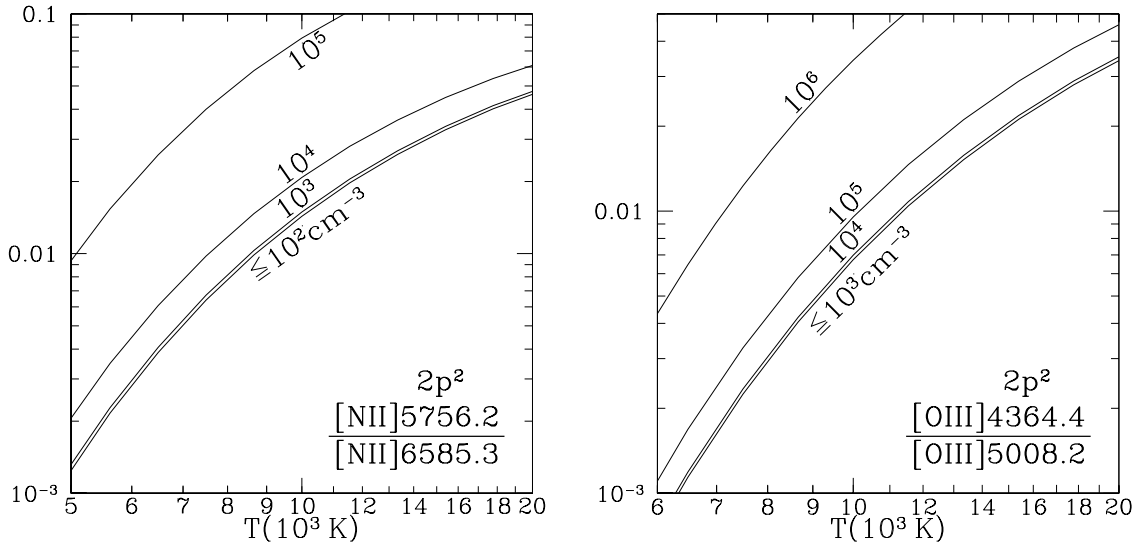


Figure 4.18: Example line ratios used for temperature diagnostics. [Draine 18.2]

4.6.8 Density Diagnostics

We can use fine-structure lines as density diagnostics. The energy difference between these levels is tiny, and therefore all collisions can cause excitation to the upper level. The density will determine the line strength ratios, not the temperature.

We want here ions with triplet ground states. Again NII and OIII are the most often used. At low densities, collisional excitation to will be followed by radiative decay. Therefore, the

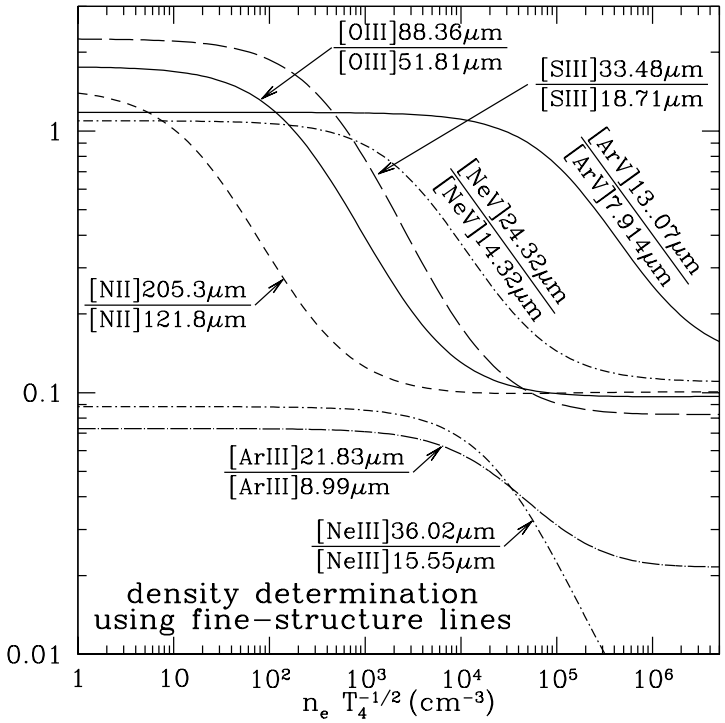


Figure 4.19: Densities using fine-structure lines.

radiated power is imply related to the collision rates:

$$\frac{j(2\to 1)}{j(1\to 0)} \approx \frac{\Omega_{20} e^{-E_{21}/kT}}{\Omega_{10} + \Omega_{20} e^{-E_{21}/kT}} \frac{E_{21}}{E_{10}} ., \tag{4.126}$$

where we have assumed $A_{20} \ll A_{10}$ to go from Equation 4.125.

In the high density limit, we can assume $T_{\rm ex} = T$ and use the Boltzmann equation. Taking $j_\nu = 1/(4\pi) n_u A_{u\ell} h\nu \phi_\nu$,

$$\frac{j(2\to 1)}{j(1\to 0)} = \frac{n_2}{n_1}\frac{A_{21}}{A_{10}}\frac{E_{21}}{E_{10}} = \frac{g_2 A_{21} E_{21}}{g_1 A_{10} E_{10}} e^{-E_{21}/kT} \approx \frac{g_2 A_{21} E_{21}}{g_1 A_{10} E_{10}}, \tag{4.127}$$

since the energy difference of these transitions is so small compared to kT.

The additional benefit of using these fine-structure lines is that they are essentially free from extinction, since they are found in the mid- to far-infrared. The bad thing is many are difficult to observe, because ground-based facilities cannot observe those wavelengths.

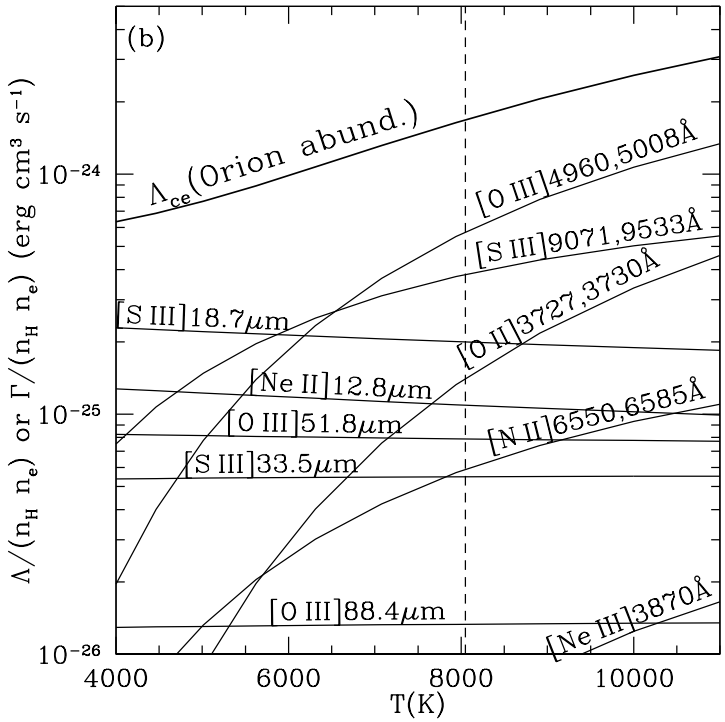


Figure 4.20: Cooling function contribution from various lines. [Draine 27.1]

4.6.9 Temperature Gradients in the Galaxy

The H II regions in our Galaxy have been found to have a temperature gradient such that H II regions near the Galactic center have low temperature and H II regions far from the Galactic center have high temperature. Why would this be?

http://adsabs.harvard.edu/abs/2011ApJ...738...27B
http://adsabs.harvard.edu/abs/1983MNRAS.204...53S
http://adsabs.harvard.edu/abs/1996ApJS..106..423A
http://adsabs.harvard.edu/abs/1997ApJ...478..190A

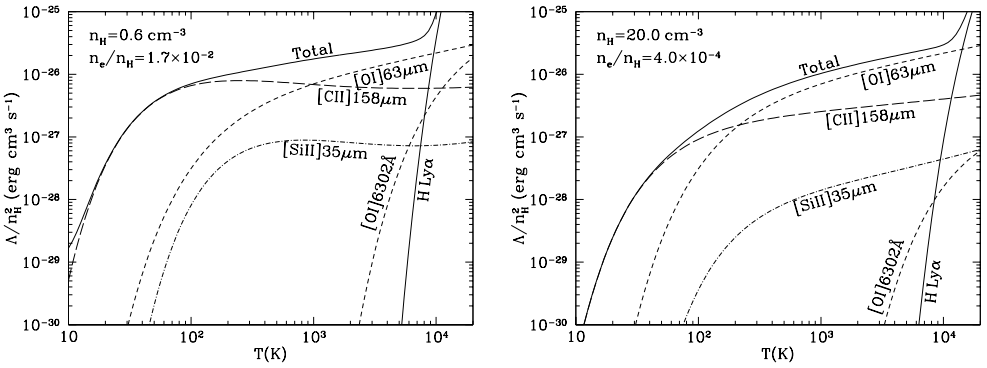


Figure 4.21: Cooling curves as a function of temperature for two densities in the ISM (from Draine book).

Chapter 5

Dust

[Pogge Notes] Up until now we have been concerned primarily with physical properties in the gaseous phases of the ISM. We now turn our attention to a solid-state component: Interstellar Dust Grains.

How important is dust? In our Galaxy the gas-to-dust mass ratio is about 100:1. Since the ISM is about 10% of the baryonic mass of the Galaxy, dust grains comprise roughly 0.1% of the total. At the same time, they absorb roughly 30-50% of the starlight emitted by the Galaxy and re-radiate it as far-infrared continuum emission. This means that only 0.1% of the baryons are ultimately responsible for a third to a half of the bolometric luminosity of the Galaxy!

Dust grains are also the primary sites of molecular formation, and are thought to be responsible for essentially all of the H_2 in the ISM. Molecular chemistry is impossible without dust grains to act as reaction sites. Finally, the formation of planetary system is thought to begin when dust grains in a protostellar disk begin to coagulate into larger grains, leading to planetesimals and eventually to planets, carrying their complex organic molecules with them. Dust is not only the principle molecule builder, it might also be thought of as one of the principal ingredients of planetary formation, and life.

Dust grains are solid, macroscopic particles composed of dielectric and refractory materials. As such, we have to deal with different and fundamentally less well-understood physics. Where before we have used quantum mechanics of atoms to explain the gas-phase spectra of atomic and molecular regions in the ISM, here we must consider macroscopic particles, and are largely dealing with the properties of solid bodies. Many of the physical details are empirical as we do not yet know the precise composition of dust grains, nor do we know their precise physical properties.

Much of the physics we will discuss is based on tentative explanations of observed phenomena. Nobody has yet convincingly been able to produce grains in the laboratory, much less reproduce the conditions they would experience in interstellar space, although great progress

is being made along these lines. Materials are known to change their properties under conditions of radiation bombardment (especially energetic particles like cosmic rays), and due to the inclusion of impurities. For example, we can measure the dielectric constants of pure water or CO_2 ices in the laboratory, but are unsure as to the degree we can rely upon those measurements for so-called "dirty ices" (those with embedded mineral impurities), or even pure ices that have been subjected to cosmic-ray bombardment in interstellar space. We can learn about dust in three main ways (Draine Ch.21 in addition to Pogge notes):

I. **Interaction with starlight:**
We infer the presence of dust grains along a given line of sight by their effects upon starlight passing through them. These effects include:
1) Total and wavelength-selective extinction of starlight passing through dusty regions due to a combination of absorption and scattering.
2) Reflection (scattering) of starlight by dusty clouds located behind bright stars (Reflection Nebulae).
3) Polarization of light either by scattering, or by passage through regions with macroscopically aligned non-spherical dust grains.
4) Absorption of starlight in Silicate bands, or various ice bands (e.g., H_2O and CO_2 ices).
5) Small angle scattering of X-Rays, resulting in "scattered halos" around X-Ray point sources.

II. **Emission from dust grains:**
Dust grains also emit electromagnetic radiation:
1) Thermal continuum emission from dust grains in radiative equilibrium with the local radiation field. This radiation emerges at mid- to far-IR wavelengths and is brightest near $100\,\mu$m.
2) Thermal continuum emission from non-equilibrium heating of tiny grains emitted at near to mid-IR wavelengths (1-25 μm). These are also sometimes known as "Sellgren Grains".
3) IR emission bands due to bending and stretching modes of heated grains.
4) Radio continuum emission from rotating grains (both electric and magnetic dipole radiation). This has only recently been discovered as part of the Galactic radio background in recent years, and the explanations are compelling but still tentative. Emission is near 30 GHz.

III. **Less Direct Methods for Learning about Dust**:
1) Gains preserved in meteorites
2) Depletion of elements in the ISM (missing elements presumably locked up in dust)
3) The abundance of H_2, which is known to form on dust grains
4) The temperature of H I and H_2, which in part is due to heating of photo-electrons ejected from dust grains.

Further Reading:
Two excellent recent reviews on interstellar dust are by Bruce Draine, the first an Annual Reviews article [2003, ARAA, 41, 241] and his Saas-Fee lectures from 2003. Both are available in PDF format from astro-ph (astro-ph/0304489 and astro-ph/0304488, respectively) and

online. They cover the same material, but the Saas-Fee lectures go into a little more depth than the ARAA article. Both go beyond the basics covered here and are excellent resources for learning more.[End Pogge Notes]

5.1 Interstellar Extinction [Draine Chapter 21]

Let's first deal with the interaction of dust with starlight. Barnard (1907, 1910) first noticed that stars were dimmed by an absorbing medium. This was confirmed by Trumpler (1930).

The interaction of light with dust remains the most direct way to study dust. Let's assume we are observing a star through a dusty region. Considering absorption only,

$$I_\lambda = I_{\lambda,0} e^{-\tau_\lambda} \,. \tag{5.1}$$

This assumes that all of the extinction is between the source and us, and that there is no scattering of light *into* our line of sight (which would introduce the source function into the above equation). τ_λ here is the dust extinction. Remember: extinction = absorption + scattering!

Rather than deal with τ_λ, astronomers prefer to work with A_λ, the extinction measured in magnitudes:

$$A_\lambda = m_\lambda - m_\lambda^0 \,, \tag{5.2}$$

where m_λ is the magnitude observed and m_λ^0 is the un-extincted magnitude that would be measured in the absence of extinction. You can see that $m_\lambda > m_\lambda(0)$, so A_λ is positive. In the situation with only extinction,

$$A_\lambda = 2.5 \log_{10}[F_\lambda^0 / F_\lambda] \tag{5.3}$$

or, plugging in Equation 5.1,

$$A_\lambda = 2.5 \log_{10}(e) \times \tau_\lambda \approx 1.086 \tau_\lambda \,. \tag{5.4}$$

Sometimes you will see

$$[F_\lambda^0 / F_\lambda] = 10^{-0.4 A_\lambda} \,. \tag{5.5}$$

[Kwok 10.2] It is also useful to work with our linear extinction coefficient κ_ν. For a spherical object of radius a, the absorption coefficient in cm^{-1} is:

$$\kappa_\nu = \pi a^2 n_d \,, \tag{5.6}$$

where n_d is the dust number density. When $a \gg \lambda$ however, this equation is not valid. We need to introduce a dimensionless efficiency parameter $Q_\nu(a)$, such that

$$\kappa_\nu = \pi a^2 Q_\nu(a) n_d \,, \tag{5.7}$$

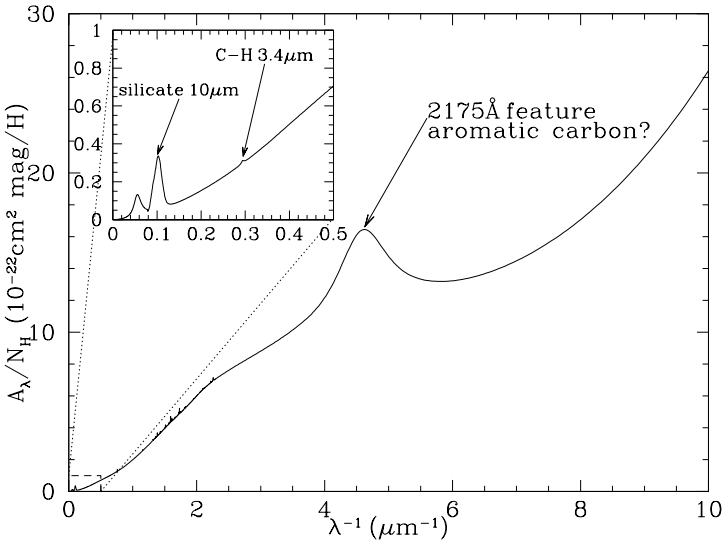


Figure 5.1: Extinction versus inverse wavelength on a typical sight line in the local diffuse ISM [Draine 21.1]

Note that this is just a different way of dealing with the cross section; we'll return to the efficiency parameter later. For gas-phase physics, we computed different cross sections for each transition. For dust, we normalize the physical cross section by the efficiency parameter.

Of course, $\tau_\nu = \int \kappa_\nu \, ds$, so

$$\tau = \int \pi a^2 Q_\nu n_d \, ds = \pi a^2 Q_\nu N_d = \frac{n_d}{n_H} \pi a^2 Q_\nu N_H \, . , \tag{5.8}$$

where N_d is the dust column density, N_H is the hydrogen column density, and the same for the volume densities. When combined with Equation 5.4,

$$A_V = 1.086 \pi a^2 Q_V N_d = 1.086 \frac{n_d}{n_H} \pi a^2 Q_V N_H \, . \tag{5.9}$$

The fraction n_d/n_H is the "dust to gas number ratio." We frequently want the gas-to-dust mass ratio, which is usually assumed to be 1/100. Be careful!

5.1.1 The Reddening "Law"

The "extinction curve," which is A_λ as a function of frequency rises with decreasing wavelength. This gives rise to "reddening" of stars as the blue light is selectively attenuated. The exact form of the reddening curve tells us about the composition and size distribution of the dust particles.

We characterize the extinction curve with a parameter R_V, the ratio of total to selective extinction:

$$R_V = \frac{A_V}{A_B - A_V} = \frac{A_V}{E(B-V)} \, , \tag{5.10}$$

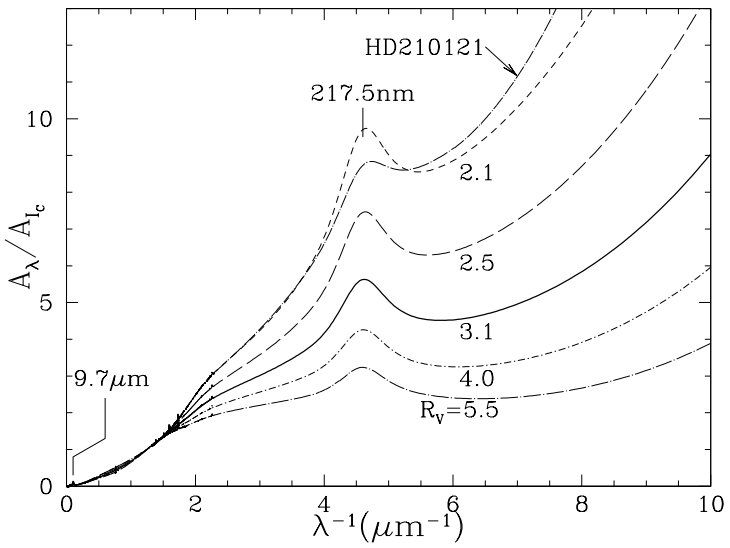


Figure 5.2: Extinction relative to extinction in band I for various values of R_V [Draine 21.2]

where V subscripts refer to the "visible" photometric band ($\lambda \approx 550\,\mathrm{nm}$), and B refers to the "blue" photometric band ($\lambda \approx 430\,\mathrm{nm}$), and $E(B-V)$ is the **color excess**, or simply the **reddening**.

The average value of R_V in the MW is 3.1. This varies, however. In dense molecular clouds $R_V \approx 5$ (perhaps due to differences in the grain size distribution). The values for R_V are not bimodal though, and many values have been measured (Draine says 2.1 to 5.7). R_V is an *empirical* factor.

Higher values of R_V lead to flatter extinction curves (less wavelength dependence). As $R_V \to \infty$, we have an ideal "gray body" absorber. That R_V is not close to ∞ tells us that dust does not absorb as a gray body (which is obvious from observations of reddened stars).

On example of a gray body is fog on Earth. There is essentially no wavelength dependence for fog. Also, when you are in fog, you are inside a cloud - how cool is that!?

5.1.2 Shape of the Extinction Curve

The extinction curve has a characteristic shape. The curve itself is usually given in A_λ/A_{ref}, where "ref" is some reference wavelength.

Draine mentions a model of Cardelli (1989) that uses six parameters to reproduce the extinction curve. The general shape of the extinction curve from the UV to NIR is $A_\lambda \propto \lambda^{-1}$ (extinction increases with decreasing wavelength). The prominent bump is called the 2175Å bump (more on this in a bit).

If the dust grains were large compared to the wavelength, the extinction cross section would be independent of wavelength, with $R_V = \infty$ (flat extinction curve). Google says fog particles

are $\sim$ 10s of microns - much larger than visible wavelengths. Since extinction rises for decreasing wavelength, the grains smaller than the wavelength must be contributing to the extinction, down to $\lambda = 0.1\,\mu$m. There must be a population of grains with radii $a \lesssim 0.015\,\mu$m.

As we have seen before for H I and molecular gas, dust is well mixed with the gas. There is a linear correlation between dust and gas quantities:

$$\frac{N_H}{E(B-V)} = 5.8 \times 10^{21} \mathrm{H\,cm^{-2}\,mag^{-1}}\,. \tag{5.11}$$

or, when $R_V = 3.1$,

$$\frac{A_V}{N_H} = \frac{3.1}{5.8 \times 10^{21} \mathrm{H\,cm^{-2}\,mag^{-1}}} = 5.3 \times 10^{-22} \mathrm{mag\,cm^2\,H^{-1}} \tag{5.12}$$

This is a pretty fantastic expression. We need only measure the extinction to derive the H I column density.

Features in the extinction curve

The 2175 Å bump The most conspicuous feature of the extinction curve is the 2175 Å bump (Figure 4 above). The central wavelength of this feature is nearly identical along all sightlines, but its width does change. It is seen in other galaxies, but the strength is correlated with galaxy metallicity. It is slightly weaker in the LMC extinction curve (metallicity $\sim$ 50% solar), but essentially absent in the SMC extinction curve (metallicity $\sim$ 10% solar). While there are many ideas, at present the carrier of the 2175 Å feature is basically unidentified. The bump "carrier" must be common in the ISM, but not necessarily present in all galaxies.

Because the feature is strong, it must be made of abundant elements: H, C, O, Mg, Si, S, or Fe. Draine says it must be C, but doesn't give us a lot more to go on here. It also must be long-lived and not easily destroyed.

[Whittet 3.5] Here is what we know observationally:
1) stars that sample the diffuse ISM tend to have relatively strong bumps
2) stars associated with H II regions have narrow bumps (Environment affects)
3) stars that sample dense clouds have broad bumps. (Environment affects)
4) the bump is generally uncorrelated with the FUV rise in extinction. From this we learn that the FUV part of the extinction curve and the bump must be caused by different carriers.

The most likely carrier is graphite of partially graphitized carbon grains. Carbon is abundant and graphite is not easily destroyed. In a lab, small carbon grains can match the location and width of the bump.

In Draine (2011), he states rather conclusively that the bump carrier is actually PAHs (see below). [Google draine PAHs review, second link]

Silicate features at 9.7 and 18 μm
The conspicuous absorption feature at 9.7 μm is due to Si-O stretching. It is seen in emission

and in absorption. The fact that the 9.7 μm band is fairly featureless, unlike what is seen in laboratory silicate crystals, suggests that this "astrophysical" silicate is primarily amorphous rather than crystalline in nature. Near 18 μm there is another feature caused by Si-O-Si bending in amorphous silicates. Whew! At least these features can be explained.

[Draine 23.4] Lending support to this identification, the 9.7 μm and 18 μm features are detected in the outflows of oxygen rich atmospheres, but not in outflows from carbon stars. Silicates should form around the oxygen rich stars, but not for carbon stars.

Laboratory attempts to recreate the silicate absorption features have shown that the material cannot be crystalline. If it were, we would have sharper features, not at all like the braod features we actually detect.

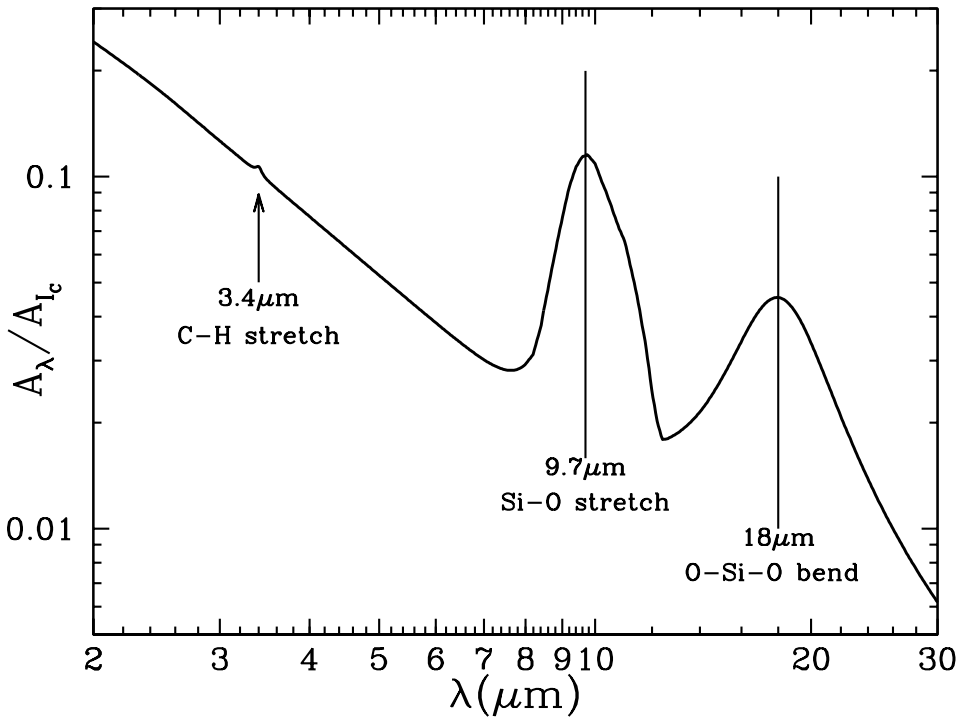


Figure 5.3: IR extinction toward the Galactic center from Keper et al (2004) [Draine Figure 23.2]

3.4 μm Aliphatic C-H feature: This is a broad extinction feature at 3.4 μm seen along lines of sight where the interstellar extinction is very high ($A_V \gtrsim 10$) associated with refractory grain material since it is often seen in regions of diffuse atomic gas. It is identified as a C-H stretching mode in "aliphatic" hydrocarbons (organic molecules with carbon atoms joined in straight or branched chains).

The origin of this feature is still unclear, but suggestions include aliphatic hydrocarbon residues produced by UV photolysis of ice mantles on grains, hydrogenated amorphous carbon, and hydrocarbon mantles on silicate grains.

Diffuse Interstellar Bands (DIBS)
Many many spectral features of the extinction curve are not yet identified. These are collectively called the "Diffuse Interstellar Bands" (DIBs). Hobbs et al (2009) report 414 DIBs between 3900 and 8100 Å. Draine is embarrassed that there is this wealth of data from the DIBS that we currently are not using. For some DIGBS, the strength of emission is correlated with that of other DIbS, one another. but for most nothing is known.

Interstellar Ices
The strongest ice feature is the 3.1 μm O-H stretch band in water (H_2O) ice, two unidentified features at 6.1 and 6.8 μm, and a band at 15.2 μm identified with CO_2 ice (see Whittet et al. 1996, A&A, 315, L375 and below). Other ice bands are CO, CH_4, NH_3, and CH_3OH. These are generally thought to arise in icy "mantles" that encase dust grains found in dense molecular clouds. Ice mantles are not found on grains in the general ISM, as exposure to the general interstellar radiation field sublimes the ices (for example, H_2O ice features in the Taurus dark cloud are only seen when $A_V \gtrsim 3.3$). The ice bands are smeared out into broad features because they are in a solid state phase condensed onto a solid grain. An ice feature called XCN at 4.62 μm is attributed to CN bonds, but the "X" carrier is so far unidentified.

CO is most commonly observed in the gas phase, but it can condense as a "frost" onto dust grains when the temperature drops below ~ 17 K. Such condensation may lead to significant depletion of CO out of the gas phase deep inside molecular clouds. CO_2 has not yet been positively observed in the gas phase (despite numerous searches) but it is seen as an ice condensed onto grains surfaces. In both cases, the shape of the ice band depends on the presence of H_2O and the state of the molecules in the ice phase.

X-Ray Absorption Edges
Dust grains can also absorb and scatter X-rays, although to an X-ray photon, a dust grain looks like a dense cloud of atomic gas, with the energies of the edges modified by being in solid materials rather than in the gas phase. Photoelectric absorption edges have been seen for C, O, Fe, Mg, and Si with Chandra and XMM.

Continuous Emission
Two continuous emission components can arise from dust:
1. The "Extended Red Emission" (ERE), a broad featureless emission band peaking between 6100 Å and 8200 Å. In some nebulae this can contribution as much as 30-50% of the flux in the photometric I band (centered at 8800 Å). It is almost certainly photoluminescence: absorption of a UV or optical photon followed by re-emission. In some nebulae the conversion efficiency can be as high as 10%. The most likely photoluminescent material is some kind of carbonaceous material, but no conclusive identification with a particular carrier (PAH, tiny silicate or carbonaceous grains, etc) has yet been made.
2. Thermal continuum from dust grains. There are two forms:
a) FIR ($> 60\,\mu$m) continuum arising from warm normal-sized grains in thermal equilibrium with the ambient radiation field ($T_d = 20 - 40$ K). "Normal" size is $> 0.01\,\mu$m (100Å). These include cooler "cirrus" emission (grains in equilibrium with the ISRF) and warmer dust associated with star clusters, esp. in star formation regions.

b) 3-30 μm continuum arising from non-equilibrium heating of tiny grains (sizes of 5-50Å) to temperature of a few hundred to a few thousand K.

In general, the thermal emission is not well described by a blackbody radiation, but rather is a blackbody spectrum modified by a wavelength-dependent emissivity. We will discuss this in detail later. [End Pogge notes]

5.1.3 Scattering and Absorption by Small Particles [Draine Chapter 22]

Cross Sections and Efficiencies

Let's first define new terms!
The absorption cross section is $C_{\rm abs}(\lambda)$
The scattering cross section is $C_{\rm sca}(\lambda)$
The extinction cross section is $C_{\rm ext}(\lambda) = C_{\rm abs}(\lambda) + C_{\rm sca}(\lambda)$
We can also define τ in terms of $C_{\rm ext}$: $\tau_\lambda = n_d C_{\rm ext} L$, where n_d is the dust number density and L is the path length.

It is convenient to normalize the cross section terms by the dust grain volumes to make yet more terms! Here, we use the effective radius of a sphere containing the same volume as the dist grains:

$$a_{\rm eff} \equiv \left(\frac{3V}{4\pi}\right)^{1/3} \tag{5.13}$$

Remember, we had this earlier in the discussion of opacity. Now, defining efficiencies

$$Q_{\rm sca} \equiv \frac{C_{\rm sca}}{\pi a_{\rm eff}^2} \tag{5.14}$$

$$Q_{\rm abs} \equiv \frac{C_{\rm abs}}{\pi a_{\rm eff}^2} \tag{5.15}$$

$$Q_{\rm ext} \equiv Q_{\rm abs} + Q_{\rm sca} \tag{5.16}$$

How can we evaluate $Q_{\rm sca}$ and $_{\rm abs}$? [from Pogge] The index of refraction m has real and imaginary parts: $m = n - ik$. Introductory physics books only deal with the real part, which is from scattering. The imaginary part deals with absorption. If the real part is large, the grain is an effective scatterer. This is the case for dielectric or icy grains. If the imaginary part is large, the grain is an effective absorber, as is the case for metallic grains.

In the limit that there is no absorption, there is no imaginary part. Therefore, $Q_{\rm abs} = 0$ and $Q_{\rm ext} = Q_{\rm sca}$

[From Whittet, page 68] For dielectrics, $k = 0$ (pure scattering) and

$$m = n \simeq c_1 + c_2 \lambda^{-2}\,, \tag{5.17}$$

where c_1 and c_2 are constants. This is an empirically derived relation called the Cauchy formula. In general $c_1 \gg c_2$ so m is only weakly dependent on λ. Ices and silicates behave approximately as dielectrics and so follow this formula.

The way in which grains interact with light is determined by their size relative to the wavelength of light. We can think of two limits:
1) The dust grains are small compared to the wavelength. In this case, the light doesn't "see" the particles.
2) The dust grains are large compared to the wavelength.
We can parameterize the size of the grains as

$$x \equiv \frac{2\pi a}{\lambda} \,. \tag{5.18}$$

Small particle/long wavelength case, $x \ll 1$ ("Electric dipole limit") (following Whittet):

We won't go through the equations, but in the limit that the dust grain is much smaller than the wavelength of light, $a \ll \lambda$, we can use the electric dipole approximation to get:

$$Q_{\rm sca} = \frac{8}{3} x^4 \left|\alpha\right|^2 \,, \tag{5.19}$$

where α is the electric polarizability of the grain,

$$\alpha = \frac{m^2 - 1}{m^2 + 2} \,. \tag{5.20}$$

[Draine uses $\epsilon = \sqrt{m}$ here for some reason. I don't like that nomenclature.] We see immediately that $Q_{\rm sca} \propto \lambda^{-4}$. This is Rayleigh scattering, which has a very strong wavelength dependence (scatters short wavelengths much more efficiently than long wavelengths, where both "short" and "long" are large compared to the grain size). This is why the sky is blue!

If there is absorption,

$$Q_{\rm abs} = -4x \, {\rm Im}(\alpha) \tag{5.21}$$

And therefore $Q_{\rm abs} \propto \lambda^{-1}$. This result comes from the fact that α is only weakly dependent on wavelength for materials that are weakly absorbing.

[Whittet pg 70] For pure dielectrics, m is real and almost constant with wavelength (see equation above, $c_1 \gg c_2$). In this case, $Q_{\rm sca} \propto \lambda^{-4}$ and $Q_{\rm abs} = 0$. More realistically, we can expect $Q_{\rm sca} \propto \lambda^{-4}$ and $Q_{\rm abs} \propto \lambda^{-1}$.

Putting this all together, we see that at wavelengths large compared to the dust grain size,

$$Q_{\rm ext} = Q_{\rm sca} + Q_{\rm abs} \propto \lambda^{-4} + \lambda^{-1} \propto \lambda^{-1} \,. \tag{5.22}$$

This is the underlying extinction curve in the UV-to-NIR part of the EM spectrum. Neat!

Large particle/short wavelength case of $x \simeq 1$ (Mie Theory)

When the dust grain sizes are comparable to the wavelength, we cannot use the electric dipole approximation. We must find a solution to Maxwelll's equations with an incident plane wave, for an object of specified size and shape, composed of material with a dielectric function ϵ or refractive index m. This was done by Mie (1908) and Debye (1909) for spherical grains, and is known as Mie theory. Although the criterion that $x \simeq 1$ seems restrictive, most scattering happens is due to this sized grain. Polarization shows us that grains are nonspherical, however, so we may not expect Mietheory to hold exactly in the ISM.

We cannot assume that $Q_{\rm abs} \simeq 0$.

The actual solution is best done on a computer because analytically it is very complex. The solution shows that the value of $Q_{\rm ext}$ depends on the refractive index (and therefore the composition since there are real and imaginary parts related to scattering and absorption). The general trend is for $Q_{\rm ext}$ to rise until it is approximately 3 to 5 when $|m-1|x \approx 2$. $Q_{\rm ext}$ has an oscillatory effect for weakly absorbing material.

For large grains, $x \to \infty$ and all solutions have $Q_{\rm ext} \to 2$. This is called the "extinction paradox". Here's why it's weird:
As $x \to \infty$, $|m-1|x \to \infty$ and the extinction cross section is **exactly** twice the geometric cross section. We would expect the extinction and geometric cross sections to be to be equal, which would mean that the grains are perfectly efficient, $\kappa_\nu = \pi a^2 n_d$. The discrepancy between expectation and reality is due diffraction around the grains, which leads to additional small angle scattering.

We can therefore define some regimes for Mie scattering:
1) $Q_{\rm abs}, Q_{\rm sca} \to 1$ and $Q_{\rm ext} \to 2$ for $x \gg 1$ ($\lambda \ll a$). This is the "geometric limit."
2) $Q_{\rm sca}$ at maximum when $x \simeq 1$ ($x \simeq 1$). Observations at wavelength λ see grains of size λ.
3) $Q_{\rm ext}, Q_{\rm sca} \to 0$ for $\lambda \gg a$ (this is where the approximations from Section 5.1.3 come into play).

5.1.4 Polarization

Polarization of starlight was discovered serendipitously in 1949. The degree of polarization was found to be correlated with the amount of reddening, indicating that dust is the polarizing agent.

The polarization percentage typically peaks near the V band, and can be described by the "Serkowski Law":

$$p(\lambda) \approx p_{\rm max} e^{-K \ln^2(\lambda/\lambda_{\rm max})}\,, \tag{5.23}$$

with $\lambda_{\rm max} \approx 500\,{\rm nm}$ and $K \approx 1.15$.

A large compilation you should know about is that of Heiles (2000). He made a catalog of starlight polarization from the literature at the time, which reveals the large-scale magnetic

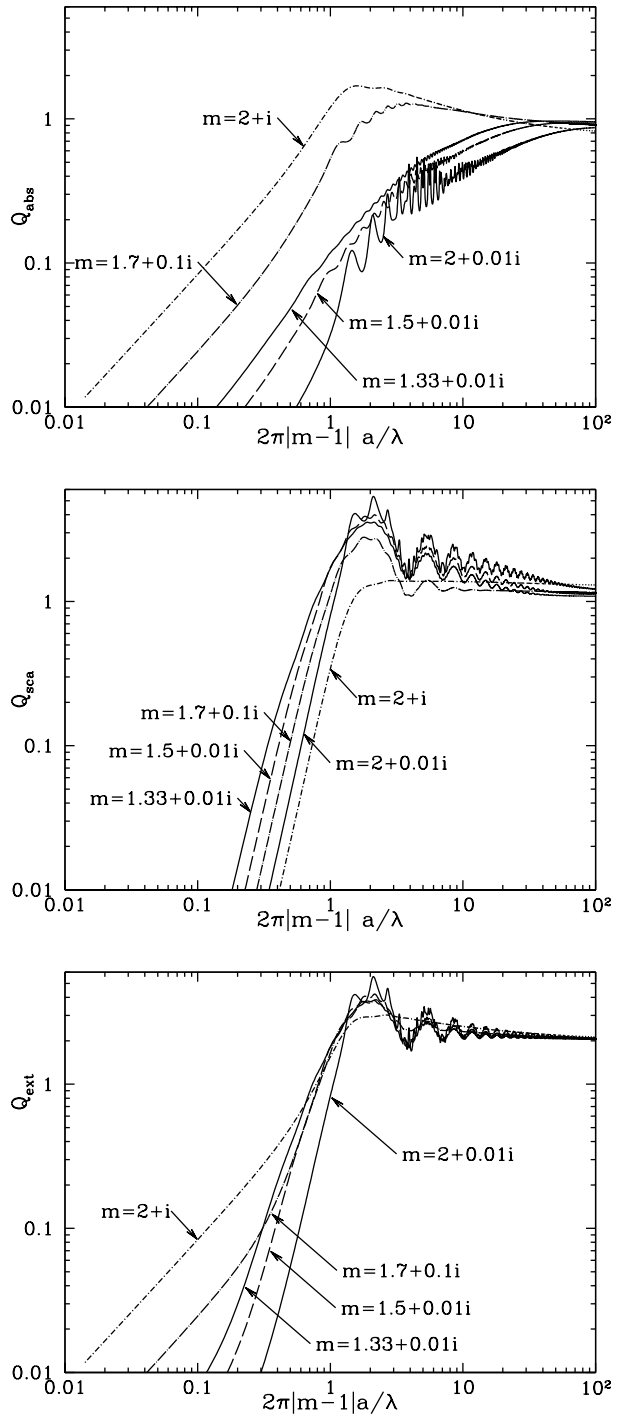


Figure 5.4: Draine figure 22.1, 22.2, and 22.3 showing $Q_{\rm abs}$, $Q_{\rm sca}$, and $Q_{\rm ext}$ from Mie theory.

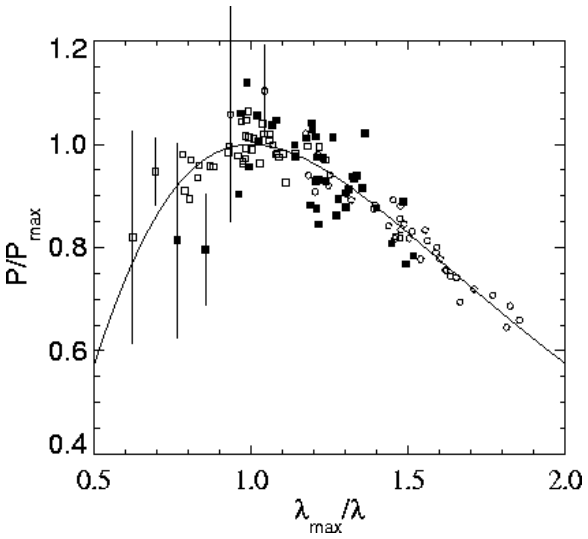


Figure 5.5: Serkowski profile fit to data, normalized by $p_{\max}$, from "UBV polarimetry of 361 A- and F-type stars in selected areas", Reiz & Franco, ApJS, 130, 13, 140. $\lambda_{\max} \approx 0.5\,\mu$m.

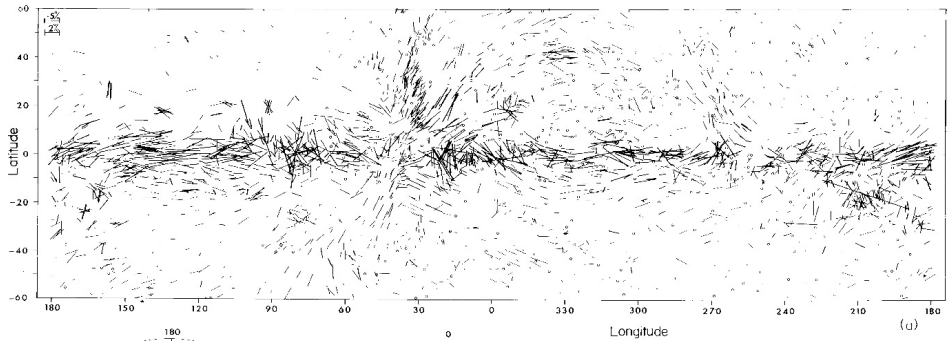


Figure 5.6: Compilation of starlight polarizations across the Galaxy showing large-scale magnetic fields.

fields in the Galaxy.

Planck also made significant progress measuring the polarization. Here is an abstract from one of their papers:
"This paper presents an overview of the polarized sky as seen by Planck HFI at 353 GHz, which is the most sensitive Planck channel for dust polarization. We construct and analyse maps of dust polarization fraction and polarization angle at 1 ° resolution, taking into account noise bias and possible systematic effects. The sensitivity of the Planck HFI polarization measurements allows for the first time a mapping of Galactic dust polarized emission on large scales, including low column density regions. We find that the maximum observed dust polarization fraction is high (pmax = 19.8%), in particular in some regions of moderate hydrogen column density ($N(H) < 2 \times 10^{21}\,\mathrm{cm}^{-2}$)..... We find that the polarization angle is ordered over extended areas of several square degrees, separated by filamentary structures of high angle dispersion function. These appear as interfaces where the sky projection of the magnetic field changes abruptly without variations in the column density.... "

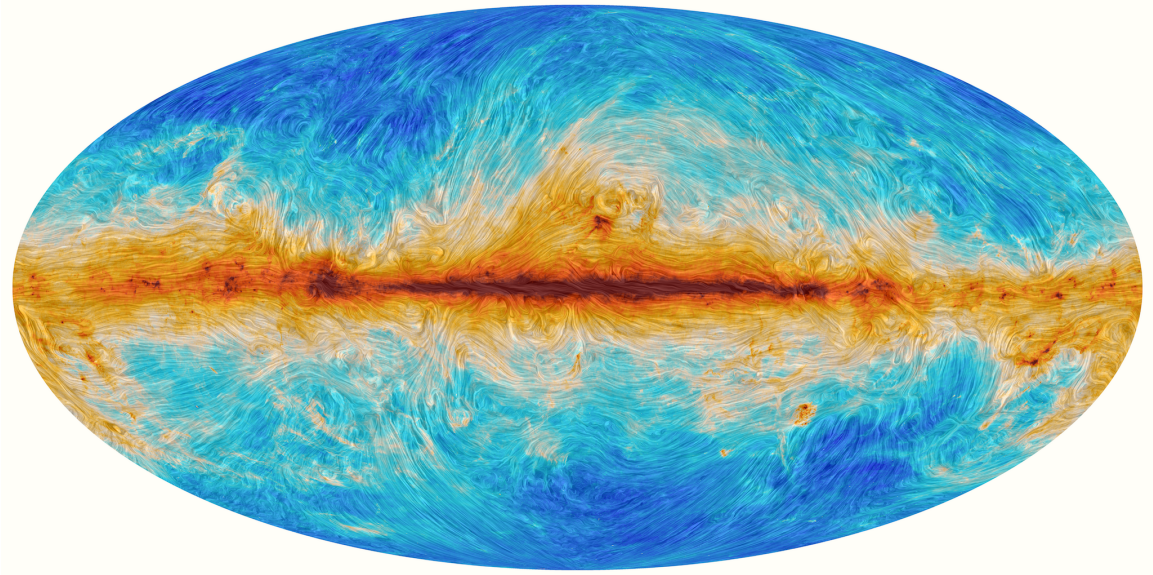


Figure 5.7: The **magnetic field** direction implied by the measured polarization, from Planck. Background is 350 GHz (850 μm) intensity.

Numerous models have been proposed for polarization caused by dust. All models require nonspherical dust grains and an alignment mechanism so that radiation sees different grain cross sections in different directions. The most promising model in my opinion is "radiative torque alignment" (RAT), first proposed by Dolginov & Mytrophanov (1976) but extended by Draine & Weingartner (1996, 1997). In the RAT model, dust grains are aligned to the local magnetic field by an anisotropic radiation field (see also Hoang & Lazarian, 2008). This theory reproduces many aspects of polarization, however fails for carbonacious grains (i.e. the 2175Åbump and the 3.4 μm feature).

5.1.5 Scattered Starlight by Dust

A reflection nebula is created when a cloud happens to be near a bright star. We see "emission" from the cloud that is actually scattered light from the nebula. We can tell this is not from the cloud itself, because the spectrum looks stellar (has absorption lines, etc.)

We can use observations of scattering and an estimate of the intensity of starlight on the cloud to estimate the albedo ω of the dust grains, and the scattering angle $\langle\cos\theta\rangle$.

The albedo

$$\omega = \frac{C_{\rm sca}}{C_{\rm abs} + C_{\rm sca}} = \frac{C_{\rm sca}}{C_{\rm ext}} \tag{5.24}$$

A pure-scattering grain has $\omega = 1$ (mirror), while pure absorbing grains have $\omega = 0$ (blackbody).

Figure 5.8: Trifid nebula at optical wavelengths. The blue emission is scattered light from hot OB stars.

The differential scattering cross section tells us about the directionality

$$\frac{dC_{\rm sca}(\theta)}{d\Omega} \tag{5.25}$$

The mean value of $\cos\theta$ for scattered light

$$\langle\cos\theta\rangle = \frac{1}{C_{\rm sca}}\int_0^\pi \cos\theta \frac{dC_{\rm sca}(\theta)}{d\Omega} 2\pi\sin\theta d\theta\,. \tag{5.26}$$

The radiation pressure cross section

$$C_{\rm pr}(\lambda) = C_{\rm abs}(\lambda) + (1 - \langle\cos\theta\rangle)C_{\rm sca}(\lambda)\,. \tag{5.27}$$

Astronomers have found that $\omega \approx 0.5$, or scattering is about as important as absorption, and $\langle\cos\theta\rangle \approx 0.5$. If $\omega = 1$, the grain would be purely scattering. This implies that the grains are somewhat forward scattering.

Rayleigh scattering for particles small compared to the wavelength has $\langle\cos\theta\rangle \approx 0$, so the particles dominating the scattering at $0.6\,\mu$m must not be small compared to the wavelength of light. They have $a \approx 0.1\,$nm.

5.2 Emission from Dust Grains

Dust grains are heated by starlight and radiate in the IR. Some dust grains are "large" and can be characterized by a single temperature. Some of the grains are "small" and are heated "stochastically".

[rest from Kwok 10.3] The densities in the ISM are frequently low, so the gas and dust are typically "decoupled." This means that the dust and gas temperature are not necessarily

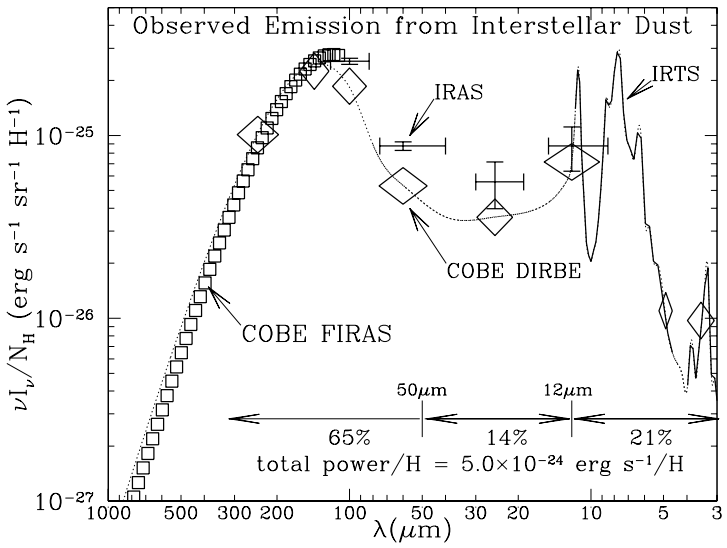


Figure 5.9: Infrared emission per H nucleon from dust [Draine 21.6].

the same. The gas temperature is the result of a balance between radiative and mechanical heating and self-radiation. The temperatures of grains are primarily determined by radiative processes: heating by diffuse starlight and cooling by self-radiation. Dust and gas temperatures can be vastly different in the diffuse ISM. In the dense ISM they can be coupled.

We can use observations of dust to derive cloud masses instead of molecular line observations. The uncertainty with molecular masses derived from CO can be large. Why? Can't do H_2 so observe CO, assume LTE ($T_{\rm ex} = T_K$), assume optically thin so ^{12}CO not great, assume $X_{\rm CO}$ and maybe ^{12}CO to ^{13}CO ratio. Assumptions with dust are different!

For observations of dust in a molecular cloud in the absence of background radiation, and if $\tau \ll 1$,

$$I_\nu = B_\nu(T_d)(1 - e^{-\tau}) \simeq B_\nu(T_d)\,\tau_\nu\,, \tag{5.28}$$

where T_d is the dust temperature. If T_d is uniform across the cloud,

$$F_\nu = \int_\Omega I_\nu d\Omega \simeq B_\nu(T_d)\Omega\tau_\nu\,. \tag{5.29}$$

After plugging in our earlier definition for τ_ν (Equation 5.8, and realizing the Q_ν earlier is actually $Q_{\rm ext}$)

$$\tau_d = \frac{n_d}{n_H}\pi a^2 Q_{\rm ext} N_H = \frac{n_d}{n_H} C_{\rm ext} N_H\,. \tag{5.30}$$

this can be rewritten:

$$N_H = \tau_d/(\frac{n_d}{n_H}C_{\rm ext}) \tag{5.31}$$

$$F_\nu = B_\nu(T_d)\Omega\tau_d \tag{5.32}$$

$$\tau_d = F_\nu/(B_\nu(T_d)\Omega) \tag{5.33}$$

$$N_H = F_\nu/(\frac{n_d}{n_H}C_{\rm ext}B_\nu(T_d)\Omega) \tag{5.34}$$

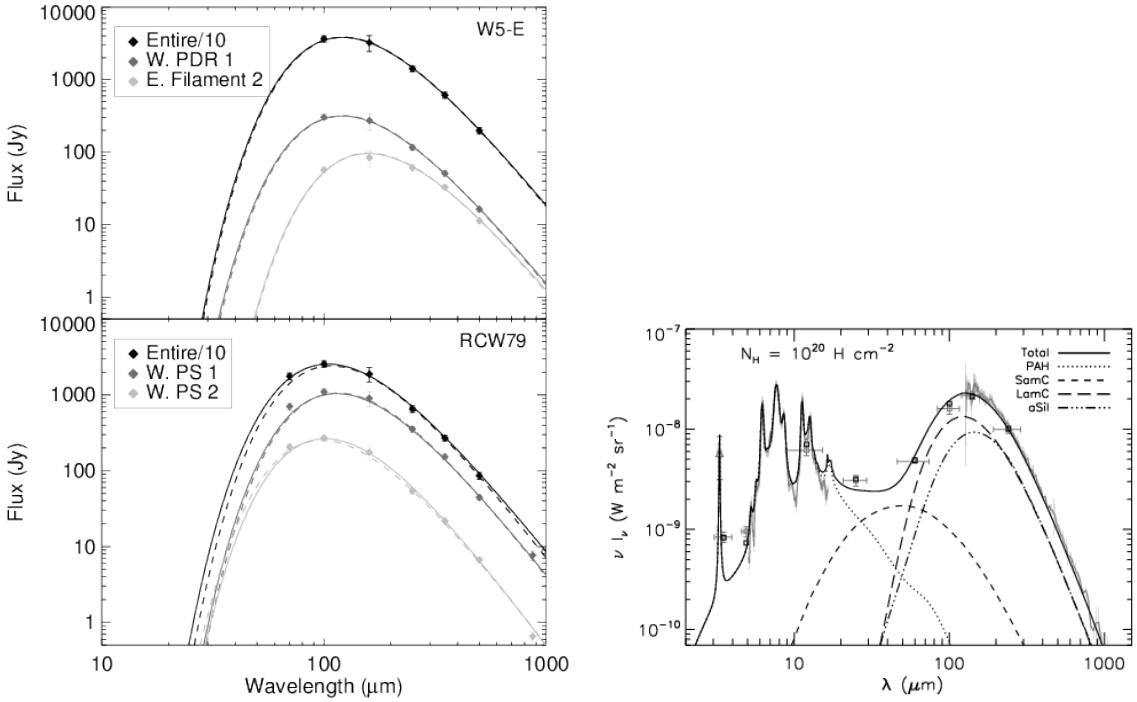


Figure 5.10: (left) Fits to SEDs from Anderson et al. (2012), showing different temperatures and column densities for three dust clumps. (right) model dust emission from the DUSTEM code (Compiegne et al., 2011).

and finally

$$N_H = R\frac{F_\nu}{2.8 m_H \kappa_\nu B_\nu(T_d)\Omega}\,, \tag{5.35}$$

where R is the gas-to-dust (mass) ratio, $R \equiv \rho_g/\rho_d$, the factor of 2.8 comes from the inclusion of helium (see HW #1) and other heavy elements, and assumes molecular material, and we have introduced a new term:

$$\kappa = \frac{n_d C_{\rm ext}}{\rho_d}\,. \tag{5.36}$$

κ here is the *mass absorption coefficient*. DO NOT GET IT CONFUSED WITH THE LINEAR ABSORPTION COEFFICIENT!!!! They are however related, since for the mass absorption coefficient $\tau_\nu = \int \kappa_\nu \rho\, ds$, with ρ the mass density. We usually parametrize κ as

$$\kappa_\nu = \kappa_0 \left(\frac{\nu}{\nu_0}\right)^\beta = \kappa_0 \left(\frac{\lambda}{\lambda_0}\right)^{-\beta}\,, \tag{5.37}$$

where β is determined through models of grains and κ_0 is some fiducial value at ν_0. In work I have done I assumed $\beta = 2$, $\kappa_{350\,\mu\rm m} = 7.3\,{\rm cm^2\,g^{-1}}$.

[How can we get temperature?] Two main ways: "color temperature" from ratio of two fluxes, or by fitting the entire SED. [How did we get temperatures for molecular gas?]

We can get cloud masses (dust+gas) from the column density (see Hildebrand (1983)):

$$M = R\frac{F_\nu D^2}{\kappa_\nu B_\nu(T_d)\Omega}\,, \tag{5.38}$$

where F_ν is the flux and D is the cloud distance. This gives cloud masses! [What are our assumptions?] All molecular (probably fine), dust to gas ratio (large uncertainties), opacity law (huge uncertainties), temperature can be fit for.

5.2.1 Observed spectral features of Dust

While dust doesn't have spectral lines per se, it does have a number of features we can use to discern its composition.

Polycyclic Aromatic Hydrocarbons

There is conspicuous emission in star-forming galaxies, predominantly at 3.3, 6.2, 7.7, 8.6, 11.3, 12.7, and 13.55 μm. These are attributed to vibrational transitions of PAH molecules. Previously called the "Unidentified IR Bands" (or UIR bands in some older papers), they are now most often referred to generically as the "PAH features." They are seen **in emission** towards PNe, H II regions, reflection nebulae, and young stellar objects, primarily in dense regions.

PAHs are large molecules or small grains, and have $\gtrsim$ 50 atoms, predominantly carbon. Their carbon is arranged into hexagonal planar rings, with H atached at the boundary. PAH emission can account for up to 20% of the total infrared luminosity of a star forming galaxy!

The 3.3 μm feature is from the C-H stretching mode (we had the 3.4 μm C-H stretching mode earlier. [Offset caused by PAH structure?]). The 6.2 and 7.7 μm features are from vibrational modes of the carbon skeleton. The 8.6 μm feature is from in-plane C-H bending, and the 11.3, 12.0, 12.7, and 13.55 μm flatures are due to out-of=-plane bending.

Real ISM PAHs will be more complicated than the simple picture. They can be ionized producing PAH^+ cations, collide with free electrons to produce PAH^- anions, have N in place of H, etc. Also, there will be a large size distribution of course.

Perhaps 10-15% of all interstellar C is in PAHs containing fewer than 500 C atoms. We know it's carbon again because planetary nebulae with high C/O ratios have stronger PAH features. This is a similar line of reasoning as we used for the 9.7 and 18 μm silicate features.

How do PAHs emit? They are small grains and so absorption of a single photon can quickly spike their temperatures. The highest energy photons will destroy PAHs though. The PAHs will absorb soft UV photons, immediately spike their temperatures to $\sim$ 1000 K, then cool via the various vibrational and bending modes. Usually called "flourescence," which is just re-emission of photons by solids on a short timescale ("phosphorescence" is a longer time scale).

While they are associated with PAH molecules, no single particular PAH "carrier" has been positively identified for any of them. The matches between observations and laboratory spectra are always close, but never exact. This has led some researchers to suspect that the PAH features arise from complex mixtures of different carriers (e.g., mixtures of neutral and positively-charged PAHs as discussed by Alamandola, Hudgins, & Sanford 1999 ApJ, 511, L115). Proposals include free PAH molecules, PAH "clusters", and particles composed at least in part of PAHs. Other researchers have investigated what happens to the spectra of various PAHs, HACs, etc. when they are "damaged" or "modified" by the harsh radiation environment of interstellar space (e.g., ionization, addition or loss of hydrogen, etc.). Since such conditions cannot be easily reproduced in the laboratory (at least, nobody has succeeded yet), this is hard to test. Detailed quantum mechanical calculations are currently beyond our computational ability for such complex molecules. Further, the origin of the PAHs remains unknown and a matter of considerable speculation.

One of the most important details about PAHs is that they flouresce in UV radiation. Therefore, they are excellent tracers of high-mass stars (which emit in the UV of course). And therefore, PAH emission has been used as a tracer of star formation itself.

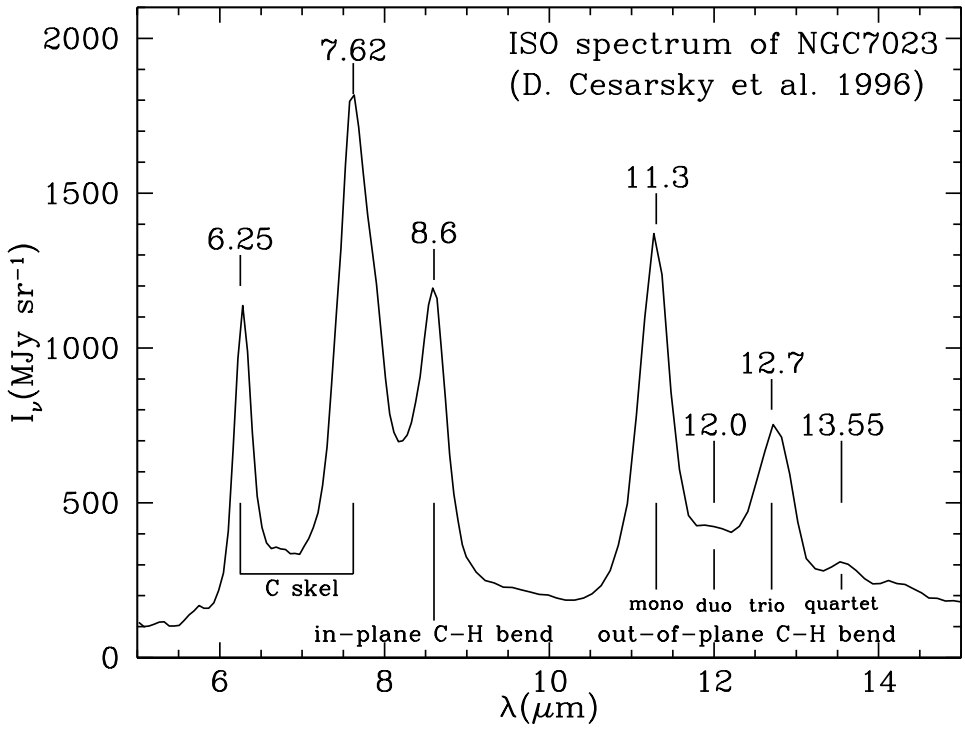


Figure 5.11: The 5 to 15 μm spectrum of the reflection nebula NGC 7023 (Cesarsky et al., 1996 [Draine 23.7]

Silicate features at 9.7 and 18 μm

See earlier!

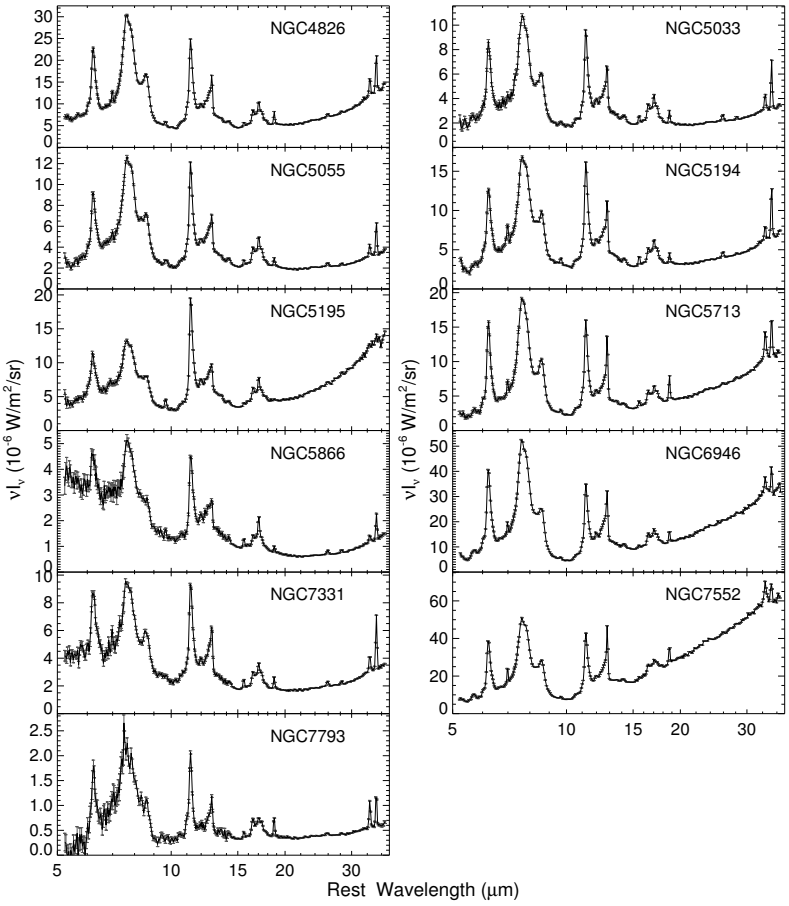


Figure 5.12: 5.5 to 36.5 μm spectra of the central regions of various galaxies. PAHs feature prominantly [Draine Figure 23.8]

Dust Emission in the 20-500 micron Range

Earlier, we reviewed the dust emission peak near 100 μm, which we ascribed to "big" dust grains. These grains are $\gtrsim$ 0.1 μm in radius. The grains are large enough to be in thermal equilibrium with their surroundings.

In massive star forming regions, there is also a second peak near 20 μm. This emission is from small grains (sometimes called "very small grains") perhaps 0.05 μm in radius. More on this the section on dust temperatures.

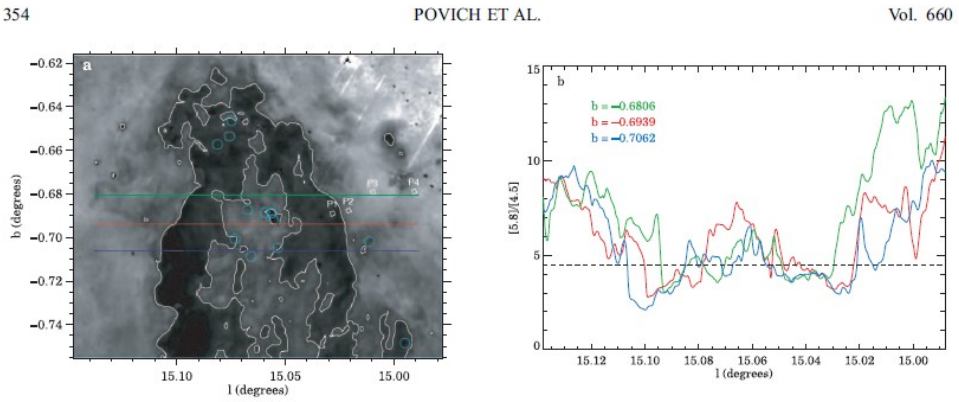


FIG. 4.—(*a*) Ratio image of IRAC [5.8]/[4.5]. The most prominent feature of this image is the dark region where PAH emission in IRAC [5.8] disappears. The boundary of this region can be traced by a fiducial contour (*white*) representing a ratio of [5.8]/[4.5] = 4.5. Four IRS SH slit pointings are overplotted, and one of them, P1, falls within the no-PAH zone. The positions of the O stars listed in Table 3 are indicated by cyan circles. Taking the band ratios in areas where the absolute mid-IR intensity is low amplifies image artifacts and noise. The bright area in the top right corner of this image takes a sight line through the molecular cloud M17 SW, where the extinction is particularly high, and hence some streaks (known artifacts of the [5.8] band) and speckling (noise) are apparent. (*b*) Image cuts in longitude *l* through the ratio image, color-coded by position in latitude *b*. All three cuts exhibit sharp dropoffs where they cross the outer boundary of the dark zone, and these transitions all span the fiducial contour value of 4.5 (*dashed line*).

Figure 5.13: From Povich et al. (2007) stdy of M17. Shows PAH destruction.

Anomalous Microwave Emission and Dust Rotational Motion

There has long been recognized to be extra emission in the $\sim 10 - 60\,\mathrm{GHz}$ regime, with a peak intensity near 30 GHz. This emission cannot be accounted for by free-free, synchrotron, thermal dust, or any other known emission mechanism. The accepted explanation is that this emission is caused by spinning dust grains. These grains have a permanent dipole moment, and therefore their rotation gives rise to emission, just like we saw for molecules.

From wikipedia: In astronomy, spinning dust is a mechanism proposed to explain anomalous microwave emission from the Milky Way. The emission could arise from the electric dipole of very rapidly spinning (10-60 G GHz) extremely small (nanometer) dust grains (Draine & Lazarian 1998), most likely polycyclic aromatic hydrocarbons. The anomalous emission was first discovered as a by-product of Cosmic Microwave Background observations which make very sensitive measurements of the microwave sky which have to identify and remove contamination from the galaxy.

Anomalous microwave emission was first seen as a surprising statistical correlation of microwave sky variations with far infrared (FIR) emission (Kogut et al. 1996, Leitch et al. 1997). This signal traced the warm galactic dust emission which was unexpected as the extrapolated infrared dust signal to microwave frequencies should have been at least an order of magnitude lower than that seen. Kogut et al. had correlated COBE Differential Microwave Radiometer observations at centimeter wavelengths with DIRBE dust emission at 140 .m, while Leitch et al. had correlated Owens Valley Radio Observatory ring observations at 14.5 and 32 GHz with IRAS 100 .m. The suggestion at the time was the correlation was due to free-free or Bremsstrahlung emission from ionized gas caused by young hot stars which are

formed in these dusty regions.

See AME in Planck data: http://arxiv.org/pdf/1309.1357v2.pdf

5.3 Dust Composition [Draine Ch. 23]

We have some evidence for the dust grain size distribution, from the extinction curve, polarization, and scattering. [How can we learn about dust composition?]

1) Spectroscopy? This is what we did for gas, but, dust spectral features are too broad and difficult to associate with any given composition.
2) Capture dust grains with a satellite. Can only do our Solar System of course, so may not be representative of the ISM as a whole, and cannot probe different environments. Number of grains captured is really low.
3) Figure out what elements are depleted in the ISM. We need a reference: the Sun or meteorites. We have to assume that the ISM has basically the same composition as the reference. This turns out to be the best method we have currently.

First we have to review abundances a little.

5.3.1 Abundances [Whittet Chapter 2]

We know that 98% of the Universe by mass is H, He. We of course are made of heavy elements. When we observe clouds in the ISM in absorption, we find that they are underabundant in certain elements with respect to the Solar abundances (measured through Solar absorption lines). This is known as **depletion**, and can be as high as a factor of 100. Depletions tell us what material is locked up in grains. Variations in depletion with environment tells us how dust grain chemistry proceeds in different environments. Let's talk about how these elements get into the dust phase.

Origins of the elements

"Condensible" elements are those that can condense from the gas phase into the dust phase.

The Big Bang made $\sim$ 90% H, $\sim$ 10% He, and $\sim$ 0% "other" by number. These are frequently denoted "X", "Y", and "Z," where "Z" refers to "metals" or to "metallicity." Present day values of Z are $\sim 2\%$.

The first generation of stars are called "Population III," and they presumably formed from material with primordial abundances. Since metals are the main coolants of the ISM, [what would this mean about their masses?] Big!!! Need more gravity to collapse.

As stars evolve, they create more and more metals. [Where does fusion take place?] In their cores, and metals are locked up there. High mass stars ($M > 8\ M_\odot$) eventually develop an iron core, with successive layers of Si, Ne, Mg, etc, like an onion. Low mass stars

($1\ M_\odot < M < 8\ M_\odot$) create elements up to oxygen. How is material returned to the ISM?
1) Mass loss from stellar winds, and
2) Supernovae (SNe), and
3) Planetary nebulae.

Stellar winds are much more massive for massive stars, but they are not enriched for massive stars [Why not?] They are not fully convective. In stars like the Sun, the convective layer goes $\sim 30\%$ of the depth. The material is locked in their cores and remains there in WD and NS. Low mass stars, however are fully convective and can return the elements made through stellar nucleosynthesis in their cores into the ISM via stellar winds. This is only up to oxygen though.

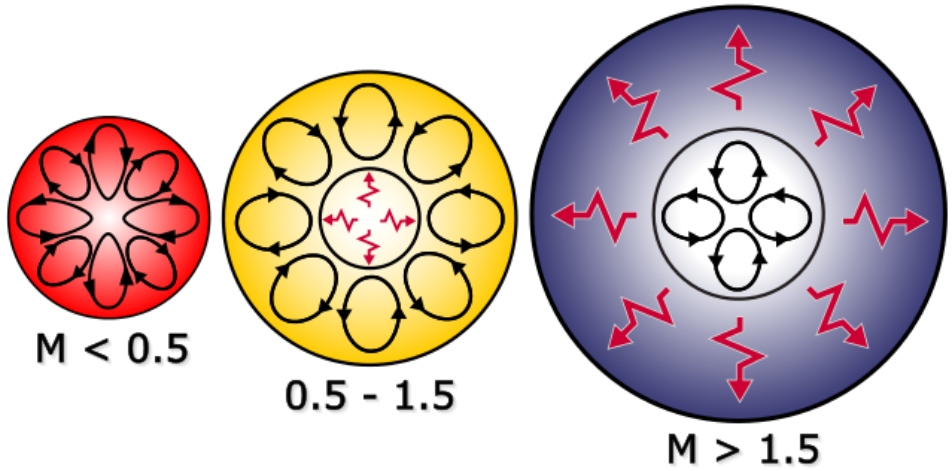


Figure 5.14: Convective and radiative zones in stars of various sizes [Wikipedia].

Low mass stars, however, do contribute significantly to the enrichment of the ISM when they evolve to become red giants, a phase called the asymptotic giant branch (AGB). In this phase the stars have an inert core of C and O, with a He-burning shell and an H burning shell. Temporal instabilities lead to photospheric enrichment by transporting C to the surface ("dredge up). O-rich stars produce silicate dust, while C-rich stars produce silicon carbide and amorphous carbon.

[See here for brief review of AGB stars: https://www.noao.edu/outreach/press/pr03/sb0307.html]

As the star evolves through the AGB phase, it cools, expands, and grows in brightness, burning its nuclear fuel faster and faster. For massive AGB stars larger than a few solar masses, the star can cool to such an extent that dust begins to condense in the outer convective envelope. At the same time, the star can begin to pulsate with very large amplitudes. As the star evolves, the pulsations become larger and longer. The large pulsation and dust formation combine to drive a wind off of the surface of the star, which can quickly lift the whole outer shell of hydrogen off, ending the hydrogen burning of the star. What is left over is a dusty shell of hydrogen slowly expanding into space, and a very hot white dwarf in the

center - objects which we know of as a bipolar or planetary nebula. The nebula disperses quickly, leaving an inert white dwarf which slowly cools.

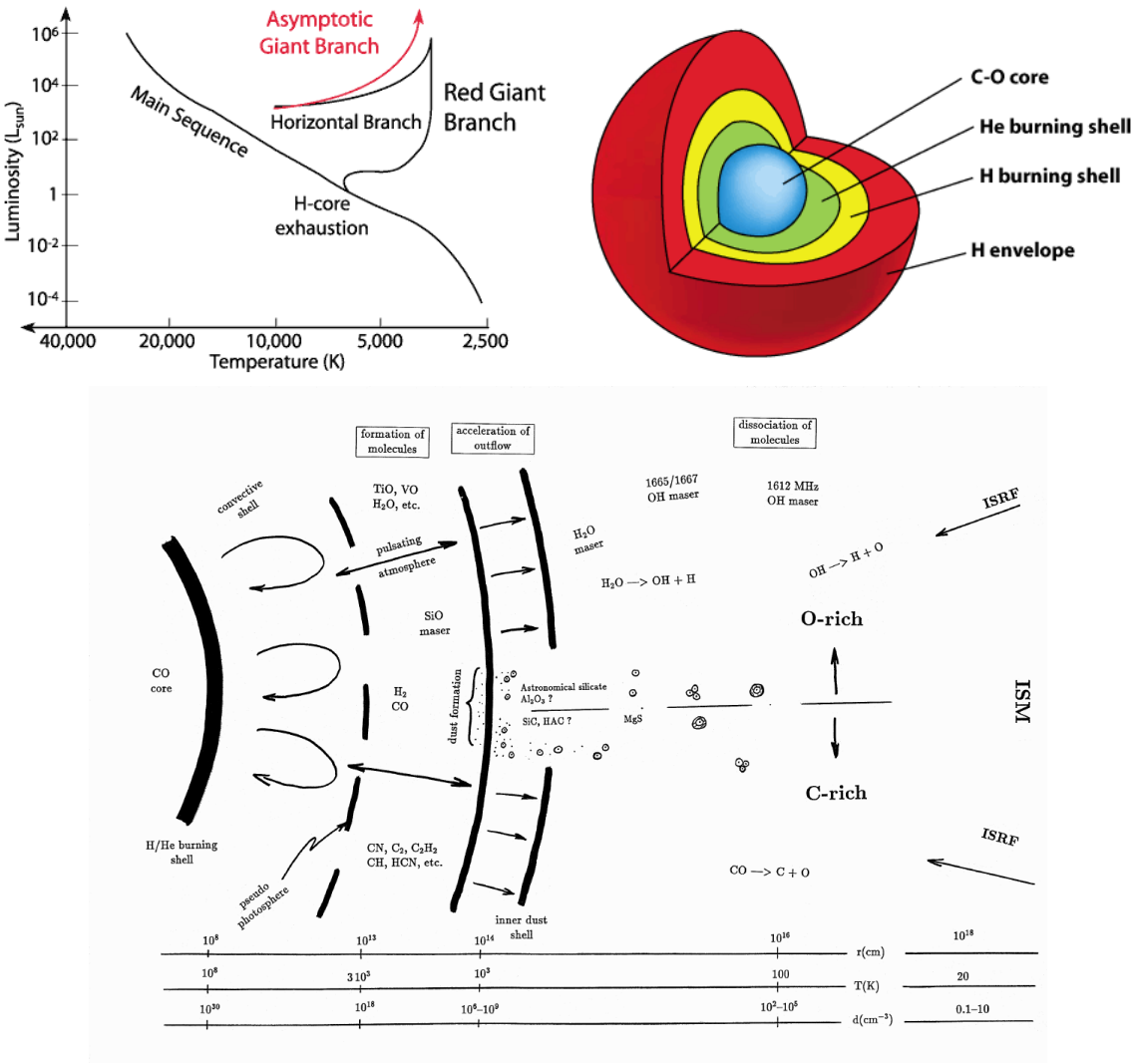


Figure 5.15: AGB stars on the H-R diagram (top left), a cutaway of their structure (top right), and a cutaway showing dust formation (bottom). [https://www.noao.edu/outreach/press/pr03/sb0307.html top two and http://aramis.obspm.fr/l̃ebertre/LNP/agb_lnp.html bottom]

SNe return elements back into the ISM. These elements can have masses up to Fe. Since massive stars evolve fast, the enrichment happens basically immediately.

PNe also return elements back to the ISM (up to O). This process can take a long time (10 Gyr) for the Sun, and most of the elements will be permanently locked up into WD.

Solar System abundances

We have two ways of measuring SS abundances: from the Sun and from meteorites. We would like to measure abundances from the Earth's crust, but "gravitational fractionation" and loss of volatiles has changed the abundances. Solar abundances should be the same as when the Sun formed. The Sun is not fully convective and therefore its photospheric composition should be unchanged. Meteorites are pristine. The ones we use for this work are called "carbonaceous chondrites;" grains have been found in these meteorites that are pre-Solar. We can measure meteoric abundances in the lab!

We find that the meteoric and Solar abundances match pretty well. Li and B are under-abundant in the Sun because their are easily destroyed. The meteoric abundances are probably more correct for these. C,N,O (and H, not shown) are under-abundant because they tend to remain in gas phase. The Solar abundances are probably more correct for these.

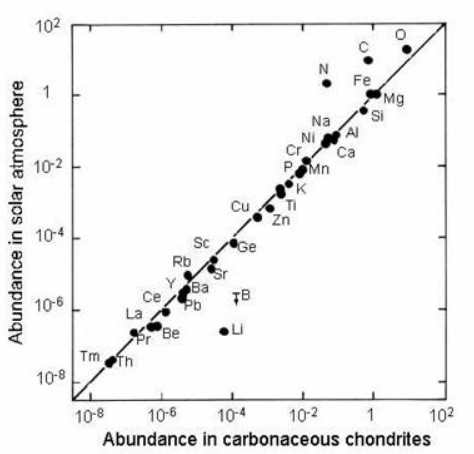


Figure 5.16: Abundances derived from the Solar photosphere versus those from carbonaceous chondrite meteorites. The line is 1:1.

The total abundances are given in the plot. The Li-group elements are easily destroyed. The regular dips are due to the fact that elements are created by the addition of α-particles (both in SNe and in stellar nucleosynthesis).

Nomenclature

Abundances can be expressed relative to $N_H = 10^{12}$. For element X,

$$\log A(X) = 12 + \log\left\{\frac{N_X}{N_H}\right\} . \tag{5.39}$$

The factor of 12 is just for conveneience since then $A(X)$ is positive even for the rarest of elements.

More commently, abundances are expressed relative to Solar:

$$\left[\frac{X}{H}\right] = \log\left\{\frac{N_X}{N_H}\right\} - \log\left\{\frac{N_X}{N_H}\right\}_\odot \tag{5.40}$$

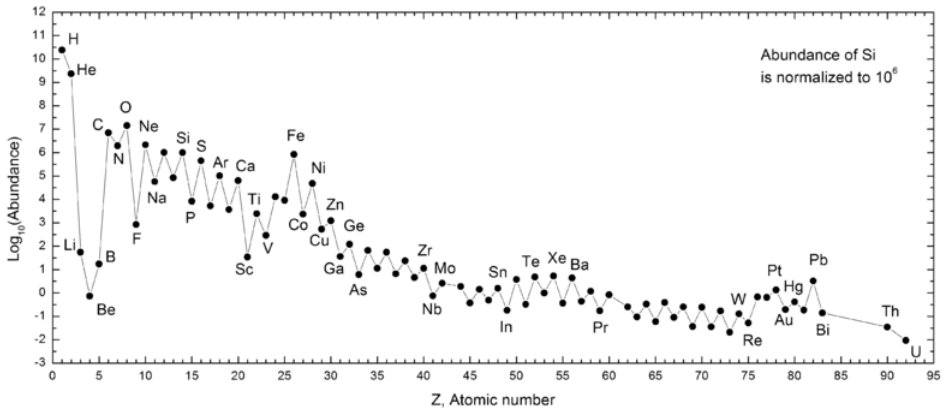


Figure 5.17: Abundances derived from the Solar photosphere.

We frequently see [Fe/H], the metallicity of a star with respect to Solar. The Sun's metallicity is slightly higher than the average for main sequence stars, 0.2 dex, where "dex" means logarithmic units. Underabundant stars with have negative values of [Fe/H] of course.

Abundance Constraints [Draine 23.1]

OK, enough background. We now have two lines of evidence: abundances from the Solar System and abundances from the gas-phase of the ISM (spectroscopy). We see below that C, Mg, Si, and Fe are underabuindany in the gas ("depleted"), with 2/3 of the C and 90% or more of Mg, Si, and Fe. These elements must be in dust grains.

The "condensation temperature" in the figure is the temperature at which 50% of the element would be incorperated into solid material in a gas of Solar abundances, at LTE at a pressure $p = 10^2\,\mathrm{dyn\,cm^{-2}}$. The condensation temperature indicates whether an element is able to form stable solid compounds in gas of Solar composition. Elements with high condensation temperatures are underabundant in the gas phase, presumably because the atoms are in solid grains.

Draine lists the following materials formed from the depleted elements:
1) Silicates, e.g., pyroxene ($Mg_xFe_{1-x}SiO_3$) or olivine ($Mg_2xFe_{2-2x}SiO_4$; $0 \le x \le 1$)
2) Oxides of Si, Mg, and Fe (e.g., SiO_2, MgO, Fe_3O_4)
3) Carbon solids (graphic, amorphous carbon, diamond)
4) Hydrocarbons (e.g., polycyclic aromatic hydrocarbons, PAHs)
5) Carbides, particularly silicon carbide (SiC) 6) Metallic Fe

There are of course other elements, but their abundances are low enough that they can be ignored.

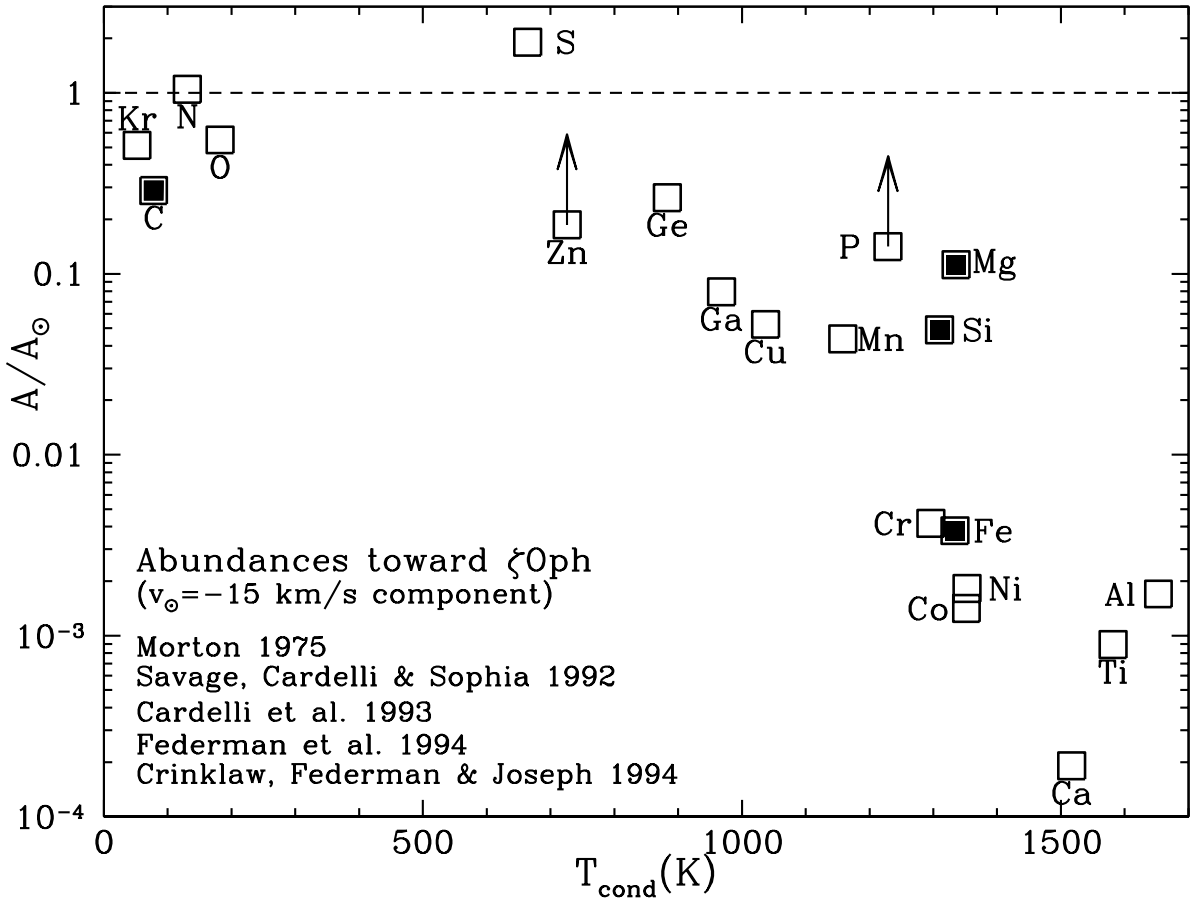


Figure 5.18: Gas-phase abundances relative to Solar in the diffuse cloud toward ζ Oph [Draine Fig. 23.1].

5.4 A Model for Interstellar Dust [Draine 23.10]

We now have all the major components of dust identified. We can finally create a model for ISM dust emission. It is worth noting that we cannot yet create a *unique* dust grain model, as there are many models that can fit the data equally well.

5.4.1 The Size Distribution of Dust Grains

The model must account for (mostly from Draine, 2011):
1) Starlight polarization: The extinction is polarization-dependent, requiring that some of the grains be nonspherical and aligned with respect to the Galactic magnetic field.
2) Scattering of starlight: A substantial fraction of the observed extinction in the optical is due to scattering, requiring that some of the grains must be large enough to efficiently scatter optical light.
3) Abundance constraints: The grain model should incorporate elements only as allowed by the total abundance of the element minus the fraction observed to be in the gas phase.

4) Infrared emission: The dust model should reproduce the IR emission spectra observed from regions with different intensities of starlight heating the dust (various "spectral line" features, black body emission near 100 μm, etc).
5) X-Ray scattering: The dust composition and size distribution must be such as to reproduce the observed strength and angular distribution of X-ray scattering by interstellar dust.
6) Microwave emission: Dust-correlated microwave emission is attributed to rotational emission from very small dust grains (Draine & Lazarian 1998a,b). This constrains the abundances of the smallest grains.
7) Dust in meteorites: Presolar grains with isotopic anomalies are found in meteorites. These were part of the interstellar grain population 4.6 Gyr ago, when the solar system formed.

The most successful models have consisted of two materials:
1) Amorphous silicates, and
2) carbonaceous material.
The carbonaceous material must also contain PAHs. We will assume these two components in the rest of our discussion of dust.

Mathis et al. showed that models with silicate and graphic spheres with a power-law size distribution $n(a) \propto a^{-3.5}$, for $a_{\rm min} < a < a_{\rm max}$ could reproduce the extinction curve. This is called the Mathis-Rumpl-Nordsieck, or MRN, model. PAHs need to be added to the MRN above model, either as a third component or as an extension of the carbonacious material to smaller sizes.

5.5 Dust Grain Temperatures [Draine Chapter 24]

We can specify the energy content of dust grains with a temperature. Draine mentions that this temperature definition is slightly problematic in the limit of small internal energies. It is also a bit confusing when discussing stochastically heated grains, because individual grain temperatures fluctuate wildly (see small grains discussion below).

Dust grains absorb energy from the ISRF, or by collisions with molecules. [Why not grain-grain collisions? Too infrequent.] In the diffuse ISM, it's the ISRF that does the heating. In dense clouds, inelastic grain-molecule collisions.

Dust grain temperatures depend on their sizes. We will break things down into "large" grains with $a \gtrsim 0.03$ μm and "small grains" with $a \lesssim 0.03$ μm. The large grains are the ones responsible for the extinction in the optical.

Just like for molecules or atoms, when grains absorb photons they usually store that energy by exciting an electron. They de-excite non-radiatively with the energy going into many vibrational modes - i.e., heat. (Some small fraction will de-ecite radiatively, or will lose electrons due to photoionization).

This is a different definition of temperature compared to atoms and molecules! For atoms

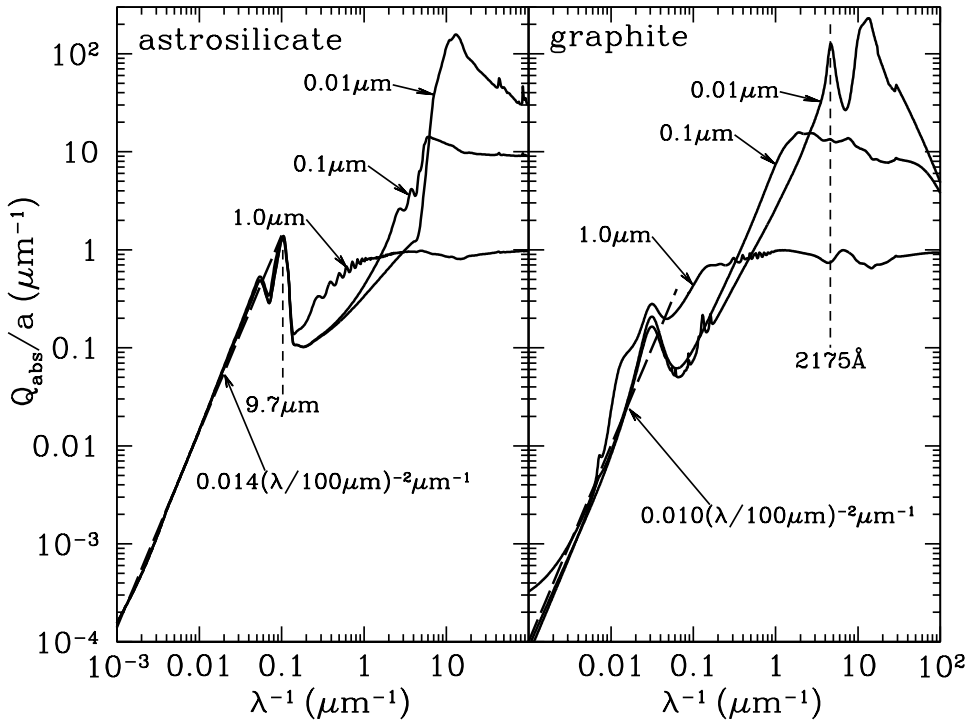


Figure 5.19: [Draine 24.1]

and molecules, the temperature is related to how fast they are moving. They have negligible heat capacity.

Ignoring radiative de-excitation and photoionization, the heating energy gain is

$$\left(\frac{dE}{dt}\right)_{\rm abs} = \int \frac{u_\nu\, d\nu}{d\nu} ch\nu Q_{\rm abs}(\nu)\, \pi a^2\,. \tag{5.41}$$

Here $u_\nu\, d\nu$ is the number density of photons with frequencies in $[\nu, \nu + d\nu]$, the photon energy is $h\nu$, and the absorption cross section is $Q_{\rm abs}(\nu)\,\pi a^2 = \kappa_\nu/n_d = C_{\rm abs}$. Different types of grains of course have different values of $Q_{\rm abs}$ [Draine Figure 24.1].

5.5.1 Radiative Heating

We can also define a **spectrum averaged absorption cross section**:

$$\langle Q_{\rm abs}\rangle_* = \frac{\int d\nu\, u_{*,\nu} Q_{\rm abs}(\nu)}{u_*}\,, \tag{5.42}$$

where $u_* = \int d\nu\, u_{*,\nu}$. The subscript here is a star because it is frequently starlight that is doing the heating. The radiative heating rate then reduces to

$$\left(\frac{dE}{dt}\right)_{\rm abs} = \langle Q_{\rm abs}\rangle_*\, \pi a^2\, u_* c\,. \tag{5.43}$$

Draine Figure 24.2 shows that $\langle Q_{\rm abs}\rangle$ is a function of radius a for gaphicte and silicate grains, and the spectrum of the ISRF. We can approximate the modeled absorption:

$$\langle Q_{\rm abs,silicates}\rangle \approx 0.18(a/0.1\,\mu{\rm m})^{0.6} \tag{5.44}$$

for silicate grains between 0.001 and 1 μmin radius and

$$\langle Q_{\rm abs,graphite}\rangle \approx 0.8(a/0.1\,\mu{\rm m})^{0.85} \tag{5.45}$$

for graphite grains between 0.005 and 0.15 μm in radius.

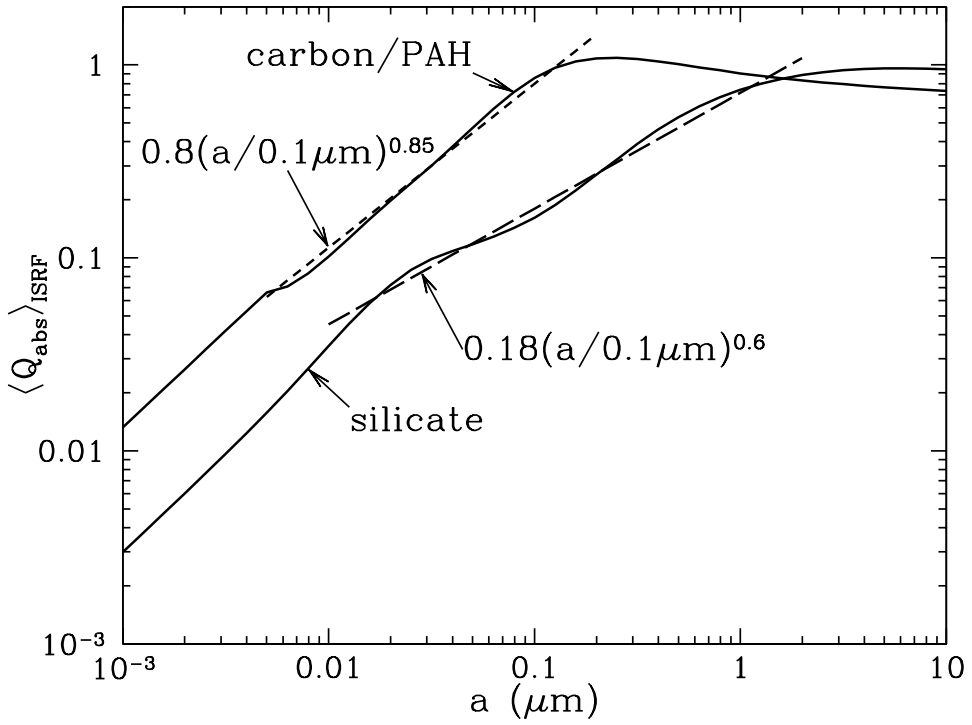


Figure 5.20: Absorption efficiency averaged over the ISFR spectrum of Mathis et al. (1983). [Draine 24.2]

5.5.2 Collisional Heating

Dust grains also can be heated via collisions with molecules. In this case Draine gives the relevant expression for spherical grains:

$$\left(\frac{dE}{dt}\right)_{\rm gas} = \sum_i n_i \left(\frac{8kT_{\rm gas}}{\pi m_i}\right)^{1/2} \pi a^2 \alpha_i 2k(T_{\rm gas} - T_d). \tag{5.46}$$

The summation is over all different gas species.
The mean kinetic energy of thermal particles is $2kT_{\rm gas}$. We would normally have $3/2kT_{\rm gas}$, but the more energetic particles collide more frequently. [Is this an exact expression then?]

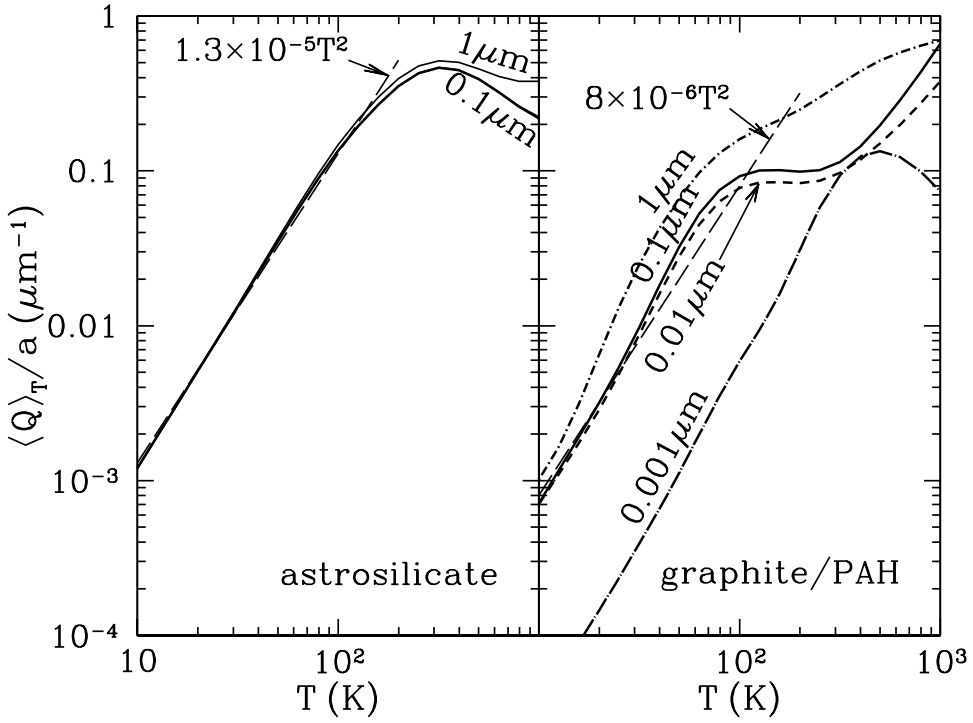


Figure 5.21: Planck averaged emission efficiency divided by grain radius as a function of grain temperature T. [Draine 24.3]

We see that if the dust and gas temperatures are the same, no energy is transferred. Makes sense!
The "accommodation coefficient" α_i ranges from 0 to 1 and measures the degree of inelasticity for collisions between particle i and a solid surface. We often do not know what value this sholud take on. [Common trick in astronomy!] Perfectly inelastic collisions have $\alpha_i = 0$ where if the molecules and grains stick together it is 1. We have to assume it is ~ 0.5 because we don't really know the grain physics well enough.

5.5.3 Collisional and Radiative heating

The ratio of the radiative and collisional heating rates tells us when one process is dominant. For hydrogen:

$$\frac{(dE/dt)_{\rm gas}}{(dE/dt)_{\rm abs}} = \frac{3.8\times10^{-6}}{U}\frac{\alpha_H}{\langle Q_{\rm abs}\rangle_*}\left(\frac{n_H}{30\ {\rm cm}^{-3}}\right)\left(\frac{T_{\rm gas}}{10^2\ {\rm K}}\right)^{3/2} . \tag{5.47}$$

For the CNM, what are typical values? $n_H \approx 30\,{\rm cm}^{-3}$, $T_{\rm gas} \approx 100$. U is the dust-weighted starlight intenisty relative to the local ISRF ($U = 1$ for the local ISFR). The ratio is $\sim 10^{-5}$ for the CNM, but much larger for dense dark clouds with smaller values of U.

5.5.4 Radiative Cooling

Dust grains cool via their infrared emission. We can think of them as blackbodies modified by $\langle Q_{\rm abs}\rangle$:

$$\left(\frac{dE}{dt}\right)_{\rm emiss.} = \int d\nu\, 4\pi B_\nu(T_d) C_{\rm abs}(\nu) = 4\pi a^2 \langle Q_{\rm abs}\rangle_{T_d} \sigma T_d^4\,, \tag{5.48}$$

where σ is the Stephan-Boltzmann constant and $\langle Q_{\rm abs}\rangle$ is the Planck-averaged eission efficiency:

$$\langle Q_{\rm abs}\rangle_{T_d} = \frac{\int d\nu\, B_\nu(T)\, Q_{\rm abs}(\nu)}{\int d\nu B_\nu(T)}\,. \tag{5.49}$$

We can approximate $Q_{\rm abs}(\nu)$:

$$Q_{\rm abs}(\nu) = Q_0(\nu/\nu_0)^\beta = Q_0(\lambda/\lambda_0)^{-\beta}\,. \tag{5.50}$$

Remember, we had this before! When talking about dust emission, we had $\kappa = \kappa_0(\nu/\nu_0)^\beta$, where κ is the mass absorption coefficint, which is linearly related to $Q_{\rm abs}$, thus the above. Draine solves the above analytically. We can get approximate relations from realizing that [Figure 24.1]

$$Q_{\rm abs,silicates} \approx 1.4\times 10^{-3}\left(\frac{a}{0.1\,\mu{\rm m}}\right)\left(\frac{\lambda}{100\,\mu{\rm m}}\right)^{-2} \tag{5.51}$$

for silicates at wavelengths $\lambda \gtrsim 20\,\mu$m, and

$$Q_{\rm abs,graphite} \approx 1.0\times 10^{-3}\left(\frac{a}{0.1\,\mu{\rm m}}\right)\left(\frac{\lambda}{100\,\mu{\rm m}}\right)^{-2} \tag{5.52}$$

for graphite at wavelengths $\lambda \gtrsim 30\,\mu$m. We can throw these into our Planck averages to get

$$\langle Q_{\rm abs,silicates}\rangle \approx 1.3\times 10^{-3}\left(\frac{a}{0.1\,\mu{\rm m}}\right)\left(\frac{T}{\rm K}\right)^{2} \tag{5.53}$$

for silicates at wavelengths $\lambda \gtrsim 20\,\mu$m, and

$$\langle Q_{\rm abs,graphite}\rangle \approx 1.0\times 10^{-3}\left(\frac{a}{0.1\,\mu{\rm m}}\right)\left(\frac{T}{\rm K}\right)^{2} \tag{5.54}$$

From Draine Figure 24.3 we see that these agree with for $T \lesssim 10^2\,{\rm K}$ (computed using actual expressions and not approximations above of course).

5.5.5 Steady State Temperature

[Group Exercise!] Compute the steady state temperature of a dust grain (planet, etc). Assume it emits isotropically. $E_{\rm out} \simeq 4\pi a^2\sigma T_{\rm ss}^4$. $E_{\rm in} \simeq L_*\pi a^2/(4\pi d^2) \simeq \pi a^2 u_* c$. [Energy density has units of energy per volume per frequency, so $u_* c$ is energy per area times time. We don't care about distance here since there are many sources of radiation.] For real grains of course, we have to factor in the absorption efficiency:

$$4\pi a^2\sigma \langle Q_{\rm abs}\rangle T_{\rm ss}^4 = \pi a^2 \langle Q_*\rangle u_* c \tag{5.55}$$

We can evaluate this equation with the solution for $Q_{\rm abs}$, assuming $Q_{\rm abs} = Q_0(\lambda/\lambda_0)^{-\beta}$. We finally get the approximation

$$T_{\rm ss,silicates} \approx 16.4(a/0.1\,\mu{\rm m})^{-1/15}U^{1/6}\ {\rm K} \tag{5.56}$$

for silicates with $0.01 \lesssim a \lesssim 1\,\mu$m and

$$T_{\rm ss,graphite} \approx 22.3(a/0.1\,\mu{\rm m})^{-1/40}U^{1/6}\ {\rm K} \tag{5.57}$$

for graphite with $0.005 \lesssim a \lesssim 0.15\,\mu$m. Remember that U is the dust-weighted starlight intenisty relative to the local ISRF ($U = 1$ for the local ISFR). Both of these are for $U \lesssim 10^4$.

We can see that larger radiation fields produce higher temperatures, although the relationship is weak. Larger grains produce lower temperatures, but the relationship is almost non-existent. We can assume that the largest grains all share the same temperature, regardless of their size distribution.

5.5.6 Temperature spikes in very small dust grains

The temperature of the grains is a measurement of their internal energies. In the case of very small grains, the absorption of a single photon will drastically change their energies (temperature). This is another way of saying that small grains have a small heat capacity. These grains do not really have a single representative temperature. This is called "stochastic heating" and is most likely the cause of the 24 μm emission seen in H II regions.

This concenpt naturally leads to the concept of a temperature distribution for each grain type, dP/dT. The distribution in temperatures is obviously larger for smaller grains. For large grains, we can actually just use a delta function centered at the steady state value.

5.5.7 IR emission from Dust grains revisited

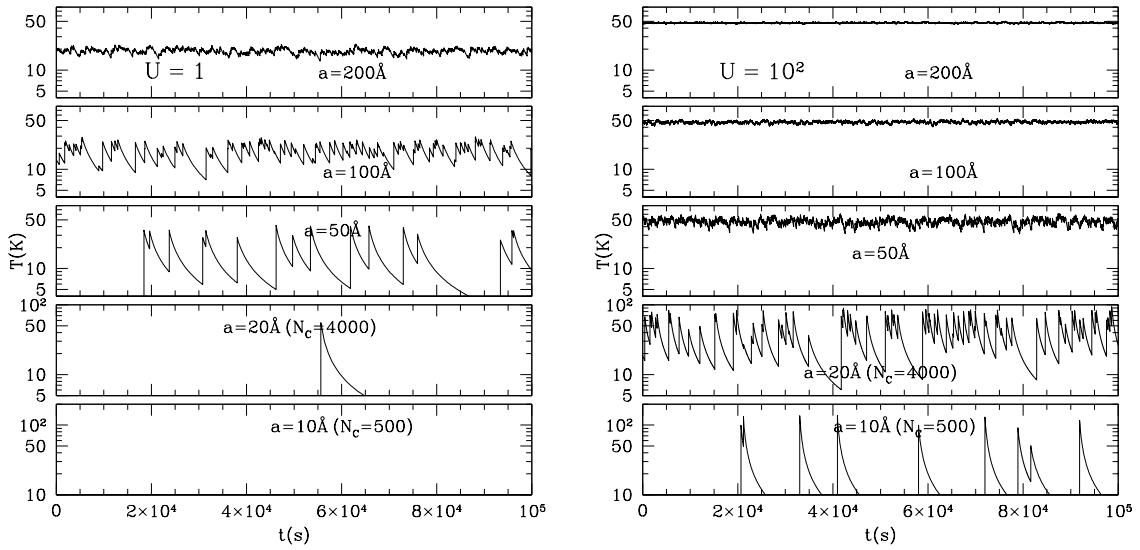
I don't know why Draine goes back to this topic, but we will follow his lead here.

In a typical spiral galaxy, perhaps 1/3 of the stellar energy is absorbed by dust grains and re-emitted in the IR. This re-emission is a quantum process, but is can be nicely approximated in a "thermal" approach.

The power per unit frequency radiated by a grain containing energy E_u is $P_u = 4\pi B_\nu(T)C_{\rm abs}(\nu)$ and the emissivity is therefore

$$j_\nu = \Sigma_i \int da\,\frac{dn_i}{da}\int dT\,\left(\frac{dP}{dT}\right)_{i,a} C_{\rm abs}(\nu;i;a)B_\nu(T) \tag{5.58}$$

Calculations of j_ν need a model for the size distribution, the temperature distribution funcion, and the absorption cross sections. Thes have been computed by Draine & Li (2007)



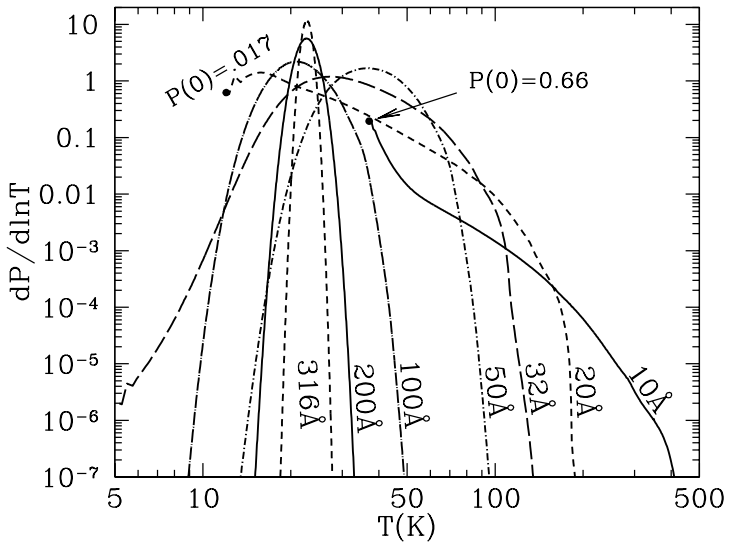


Figure 5.22: (left, right top) Temperature for carbonaceous grains. Fluctuations larger for smaller grains. [Draine 24.5] (bottom) Temperature distribution function for carbonaceous grains in ISRF with $U = 1$. P(0) is fraction in the ground state. [Draine 24.6]

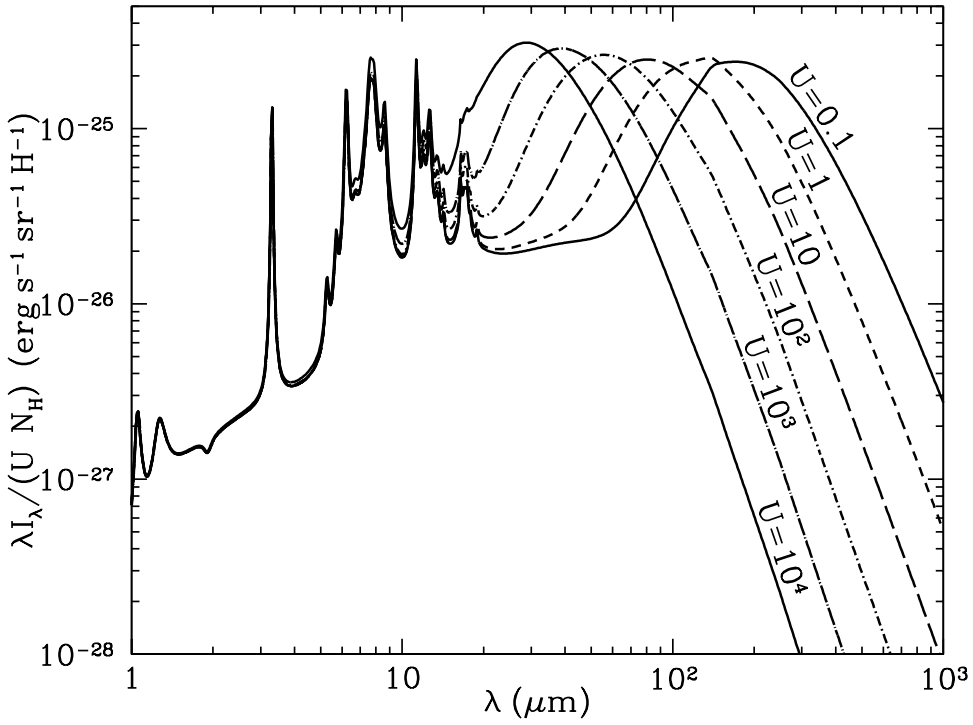


Figure 5.23: Spectra for model with silicates and graphites in various ISRFs, normalized by U (from Draine & Li, 2007) [Draine 24.7]

The PAH emission is relatively independent of the ISRF. PAHs cause more frequent temperature spikes, but the grains can cool efficiently between photons. Total energy output of PAHs is increased, but spectrum shape is not.

The large grains do become warmer as the ISRF becomes larger.

5.6 Dust Dynamics [Draine Chapter 26]

5.6.1 Radiation Pressure and Gravity [Whittet 7.3.1]

Radiation pressure can drive grains away from a star, while gravity works to keep them in orbit. The outward force due to radiation pressure is

$$F_{\rm Rad} = \pi a^2 \left\langle Q_{\rm Rad} \right\rangle \left(\frac{L_*}{4\pi d^2 c} \right) , \tag{5.59}$$

This makes a little sense, since the last terms are just the radiation pressure at distance d, and pressure is just force/area. Here $\langle Q_{\rm Rad} \rangle$ is the average value with respect to wavelength.

The gravitaional force is just

$$F_{\rm gr} = \frac{G M_* m_d}{d^2} . \tag{5.60}$$

Both of these have a d^2 dependence, due to the inverse square law. When are they equal?

$$\frac{F_{\rm Rad}}{F_{\rm gr}} = \frac{a^2 \left\langle Q_{\rm Rad} \right\rangle \left(\frac{L_*}{4c} \right)}{G M_* m_d} = \frac{3 L_* \left\langle Q_{\rm Rad} \right\rangle}{16 \pi G M_* c a s} , \tag{5.61}$$

where we substituted the specific density s of the grain material with $s = 3m_d/(4\pi a^3)$. Evaluation of this shows that graphic is easily blown away from the star, while silicates are less so. There is no dependence on this ratio with distance from the star.

This is important for AGB stars. We said earlier that dust grains are created by AGB stars. This is how they are transported away from the photosphere, into the ISM.

5.6.2 Poynting-Robertson Effect

Consider a particle in a circular orbit with radius R and velocity $v_{\rm orb}$ around a star with luminosity L_*. Because of the aberration of starlight[1], in the instantaneous rest frame of

[1]From Wikipedia: http://en.wikipedia.org/wiki/Aberration_of_light : "Aberration may be explained as the difference in angle of a beam of light in different inertial frames of reference. A common analogy is to the apparent direction of falling rain: If rain is falling vertically in the frame of reference of a person standing still, then to a person moving forwards the rain will appear to arrive at an angle, requiring the moving observer to tilt their umbrella forwards. The faster the observer moves, the more tilt is needed. The net effect is that light rays striking the moving observer from the sides in a stationary frame will come angled from ahead in the moving observer's frame. This effect is sometimes called the "searchlight" or "headlight" effect.

the orbiting paricle, the radiative flux from the star has a component $\beta L_*/(4\pi R^2)$ in the direction antiparallel to the motion of the grain, where $\beta = v_{\rm orb}/c$. This radiation therefore acts to reduce the orbital angular momentum J of the particle, causing the particle to spiral in toward the star. This is called the **Poynting-Robertson effect**.

Draine works out the magnitude of this effect. In short, the grains experience a net loss of angular angular momentum and the timescale for the orbit to decay is

$$\tau_{\rm PR} = \frac{J}{-(dJ/dt)_{\rm PR}} = 8.3\times10^7{\rm yr}\left(\frac{\rho}{3\,{\rm g\ cm^{-3}}}\right)\left(\frac{a}{\rm cm}\right)\left(\frac{R}{\rm AU}\right)\frac{1}{\langle Q_{\rm PR}\rangle_*}\frac{L_\odot}{L_*}\,, \tag{5.62}$$

where $\langle Q_{\rm PR}\rangle$ is the radiation pressire efficiency factor averaged over the stellar spectrum. We expect $Q_{\rm PR}\approx 1$ for $x = 2\pi a/\lambda \gtrsim 1$. Draine mentions that the PR effect leads to orbital decay on $\lesssim$ Gyr time scales for particles up to $\sim 10\,$cm in size. Micron-sized particles have short orbital lifetimes near stars.

Wikipedia lists the force due to the PR effect as

$$F_{\rm PR} = \frac{a^2 L_\odot}{4c^2}\sqrt{\frac{GM_\odot}{d^5}}\,, \tag{5.63}$$

which can easily be parameterized for other stars. It seems like this also contains the gravitational force?

The combined effect of these three forces (gravity, radiation pressure, and the PR effect) is that dust grains smaller than $0.1\,\mu$m are pushed out of the Solar System, and those larger tahn $0.1\,\mu$m spiral in toward the Sun. These values will be slightly different depending on the mass of the dust grain and the luminosity of the star, but the basics will be the same.

5.7 Review

We can review our knowledge of dust in a few figures:

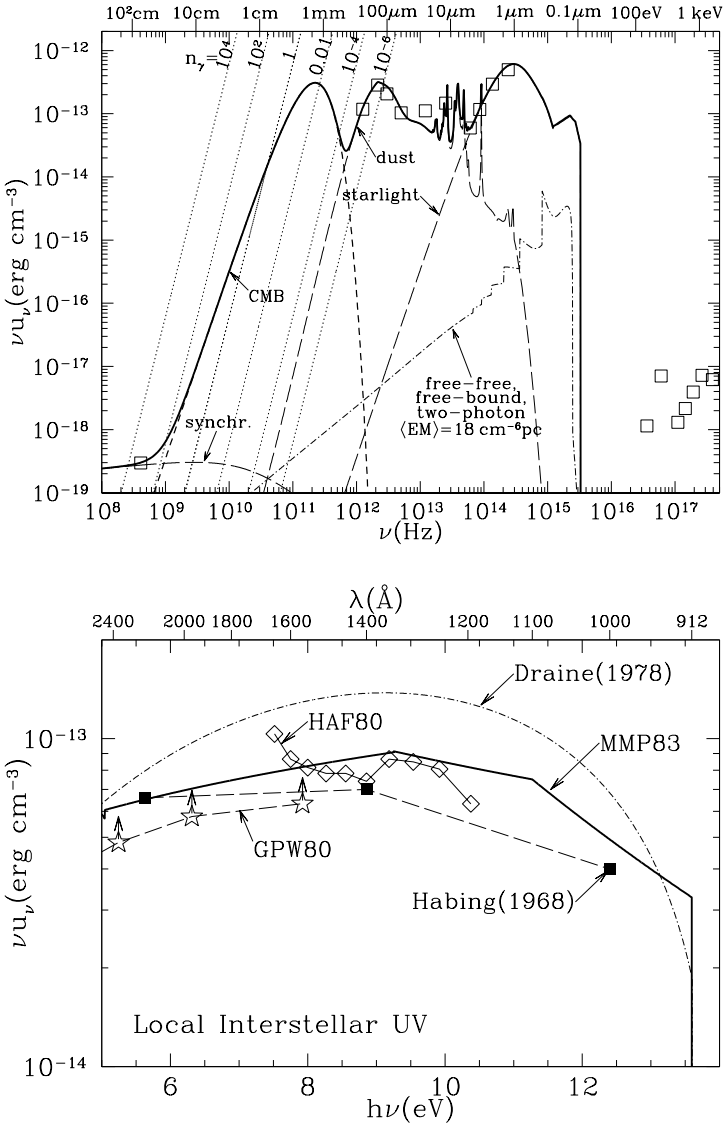


Figure 5.24: (Left) ISRF near the Sun (Draine 12.1). Spectral lines not included. Squares show measurements. Dotted lines show constant photon occupation number n_γ. (Right) The same, for a harder radiation field [Draine 12.2].

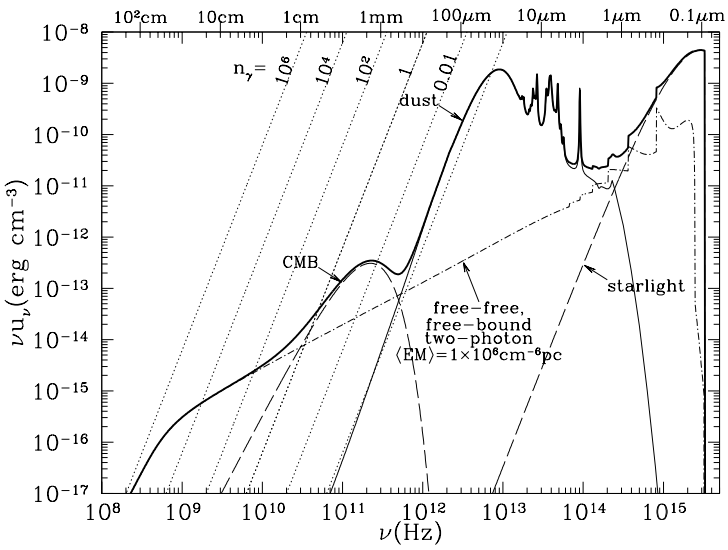


Figure 5.25: Same as Figure 5.24 (Draine 12.3), but near an O star with $T = 3.5 \times 10^4\,\mathrm{K}$. Again, spectral lines not shown and dotted lines are constant n_γ.

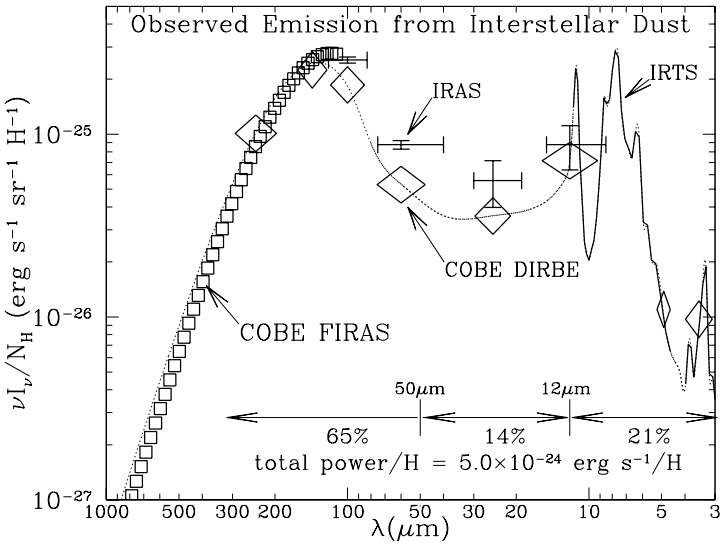


Figure 5.26: Infrared emission per H nucleon from dust [Draine 21.6].

Chapter 6

Dynamics

6.1 Shocks! (Draine Ch. 35+36) and Magnetic Fields

Shocks are created when a flow moves faster than the local sound speed. What can do this?
–novae and supernovae
–fast stellar winds
–expanding H II regions
–gas falling into the potential of spiral arms
–colliding interstellar clouds

6.1.1 The Sound Speed

What is the sound speed? In an unmagnetized gas,

$$c_s = \left(\frac{\gamma P}{\rho}\right)^{0.5}, \tag{6.1}$$

where γ is the adiabatic index. γ takes values of $\gamma = 5/3$ for ideal monotonic gas, $\gamma = 7/5$ for diatomic gas. Partially ionized gas can take values between these two extremes, but fully ionized gas has $\gamma = 5/3$. We almost always assume an ideal gas, so $P = nkT$ and therefore

$$c_s = \left(\frac{\gamma kT}{\mu m_H}\right)^{0.5}, \tag{6.2}$$

where μ is again the mean particle mass. Let's compute c_s for the various phases of the ISM!

For magnetized gas, we add in the magnetic pressure $P_B = B^2/4\pi$ so

$$c_{ms} = \left(\frac{\gamma P}{\rho} + \frac{B^2}{4\pi\rho}\right)^{0.5}, \tag{6.3}$$

where c_{ms} is the magnetosonic speed.

The ratio of the speed of the flow to the speed of sound in the gas is the Mach number in honor of Ernst Mach, a late 19th century physicist who studied gas dynamics. "Subsonic" conditions occur for Mach numbers less than one, $M < 1$. As the speed of the object approaches the speed of sound, the Mach number is nearly equal to one, $M \simeq 1$, and the flow is said to be "transonic." Supersonic conditions occur for Mach numbers greater than one, $M > 1$. Sometimes, you may hear "hypersonic," which is $M > 5$.

A shock separates $M = 0$ (the ambient medium) from $M > 1$. The shock dissipates heat, and therefore due to the entropy generation it is irreversible.

6.2 The Fluid Equations

The ISM is a fluid, and so we must treat its physics using the fluid equations. Fluid mechanics in the ISM have three main useful equations: the continuity equation, the momentum equation, and the energy equation. These are collectively known as the "conservation equations," or sometimes simply as the "fluid equations." The names clue us in to the quantity that is conserved. We will eventually add in a fourth equation from Maxwell's Laws.

6.2.1 Mass Conservation

The (mass) continuity equation conserves mass. Assume we have some comoving volume $\Omega(t)$ (apologies, I'm following Draine's notation here). From mass conservation, $\rho\Omega =$ constant. Therefore,

$$\frac{\partial}{\partial t}(\rho\Omega) = \rho\frac{\partial\Omega}{\partial t} + \Omega\frac{\partial\rho}{\partial t} = 0\,. \tag{6.4}$$

We can solve this to find

$$\frac{\partial\rho}{\partial t} + \nabla\cdot(\rho\vec{v}) = 0\,. \tag{6.5}$$

This equation is telling us that the change in density with time (the first term) must be balanced by the divergence in the quantity $\rho\vec{v}$, which is the mass flux. Any change in density must be balanced by a changing mass flow rate - either changing the flow speed or the flow density.

6.2.2 Conservation of Momentum

The momentum equation conserves momentum (duh!), and is another way of stating $F = ma$. Again, let's consider only one dimension. What forces are acting on the fluid? First, there are "body forces" that act at a distance: the electric, magnetic, and gravitational fields. Second there are "surface forces:" the pressure acting on the surface and the shear force.

We can write $m\vec{a}$ as $\rho\Omega\frac{\partial}{\partial t}\vec{v}$ and then equate that to all forces in the system:

$$\rho\Omega\frac{\partial}{\partial t}\vec{v} = \vec{F}_{\rm pressure} + \vec{F}_{\rm EM} + \vec{F}_{\rm gravity} + \vec{F}_{\rm viscosity} \tag{6.6}$$

If we have some surface element $d\vec{S}$ outward. Then, the external fluid presses inward so that the net pressure on the fluid is

$$\vec{F}_{\rm pressure} = \int(-PdS) = \int -\nabla P d\Omega\,, \tag{6.7}$$

and by Gauss's theorem,

$$\vec{F}_{\rm pressure} = -\Omega\nabla P\,. \tag{6.8}$$

For gravity,

$$\vec{F}_{\rm gravity} = (\rho\Omega)(-\nabla\Phi_{\rm gravity}) \ = \rho\Omega g,\,. \tag{6.9}$$

This is Poisson's law, albeit in a form that is probably a bit unfamiliar.

Putting it all together, after the derivation of $\vec{F}_{\rm EM}$ (which is a little involved) and the viscosity (a little confusing, but we'll ignore it later anyway), we find

$$\rho\frac{D\vec{v}}{Dt} = -\nabla\left(P + \frac{B^2}{8\pi}\right) + \frac{1}{4\pi}(\vec{B}\cdot\nabla)\vec{B} - \rho g + \hat{x}_i\frac{\partial}{\partial x_j}\sigma_{ij}\,, \tag{6.10}$$

where the term $\hat{x}_i\frac{\partial}{\partial x_j}\sigma_{ij}$ is the viscosity, which we will ignore.

A simplified form of Equation 6.10 is the case when $\vec{B} = 0$ and $g = 0$. This is then known as a "Navier-Stokes" equation.

6.2.3 Conservation of Energy

Conservation of energy is more complicated. The mechanical power or mechanical work (dE/dt) is just the pressure times the change in volume. The change in volume is just $dV = dSv$. Thus, we can integrate the momentum equation over the surface times v to find the mechanical work.

There is also heating work, which is the difference of the heating Γ and the cooling Λ.

Draine derives the expression, but it's ugly and doesn't need to be reproduced here.

6.3 The Rankine-Huginot "Jump" Conditions

Now that we have our fluid equations, we can look at the physics of shocks. We are going to be in the frame of the shock. Although the shock is propagating into the ISM, the shock frame is stationary.

If there is a discontinuity (a shock) the mass, momentum, and energy must be conserved. We can therefore set up pre- and post-sock conditions, called the "Rankine-Huginot" conditions.

We are going to be making two simplifications: first, that the flow is "steady," so that $\partial/\partial t = 0$; second, that the shock is "plane-parallel" so that the flow is entirely in $\hat{x}$ ($\partial/\partial y = \partial/\partial z = 0$). Finally, we further assume that there is a "single species," so the flow of all particles has the same velocity.

6.3.1 Conservation of Mass

If the flow is constant or "steady," $\partial\rho/\partial t = 0$. If the flow is in one dimension, we can then write:

$$\rho\vec{v} = \text{constant}\,. \tag{6.11}$$

The mass flux is a conserved quantity.

6.3.2 Conservation of Momentum

We can integrate Equation 6.10 and simplify it using the mass conservation equation and ignoring viscous forces arrive at

$$\rho v_x^2 + P + \frac{B_y^2 + B_z^2}{8\pi} = \text{constant}\,, \tag{6.12}$$

where ρv^2 is the "ram pressure," or the pressure exerted by a flow, and $\frac{B_y^2+B_z^2}{8\pi} = \frac{1}{8\pi}B_\perp^2$ is the magnetic pressure perpendicular to the flow.

6.3.3 Conservation of Energy

Ignoring viscosity again, the energy conservation equation reduces to

$$\left[\frac{\rho v^2}{2} + \frac{\gamma P}{(\gamma-1)}\right] v_x + \frac{B_y^2 + B_z^2}{8\pi} v_x - \frac{B_y B_x + B_z B_x}{4\pi} v_x - \kappa\frac{dT}{dx} = \text{constant}\,, \tag{6.13}$$

where $\kappa\frac{dT}{dx}$ refers to the heating and cooling.

In the case of $B_x = 0$ and $\kappa\frac{dT}{dx} = 0$ (no thermal conductivity), we find

$$\frac{\rho v^3}{2} + \frac{\gamma P}{(\gamma-1)} v + \frac{v B^2}{8\pi} = \text{constant}\,. \tag{6.14}$$

6.3.4 Conservation of Magnetic Flux

Although not from a fluid equation, we must also conserve magnetic flux. Maxwell tells us for infinite electrical conductance:

$$\frac{\partial\vec{B}}{\partial t} = \nabla\times(\vec{v}\times\vec{B})\,. \tag{6.15}$$

If $\partial/\partial t = \partial/\partial x = \partial/\partial z = 0$, and $B_x = 0\,$,

$$vB = \text{constant} \tag{6.16}$$

6.3.5 Solutions to the Jump Conditions

We can now put everything together to determine how our variables of interest change post-shock. The quantity on the left hand side of the fluid equations must be the same pre and post shock, so for instance a change in density must be compensated by a change in some combination of velocity, pressure, or magnetic field strength. Let's assume that we know or can estimate the pre-shock quantities $\rho_1, P_1, v_1,$ and B_1, where "1" refers to pre-shock gas and we will use "2" for post-shock gas. We then have four equations and four unknowns.

We will follow the usual convention and replace v_x with u, so in the shock reference frame $v_s = u_1$. We can also define $x \equiv \rho_2/\rho_1$, so therefore from mass conservation $u_2 = v_s/x$ and from magnetic flux conservation $B_2 = xB_1$. The momentum and energy equations are then

$$\rho_1 v_2^2 + P_1 + \frac{B_1^2}{8\pi} = \frac{\rho_1 v_2^2}{x} + P_2 + \frac{B_1^2}{8\pi}x^2 \tag{6.17}$$

$$\frac{1}{2}\rho_1 v_s^3 + \frac{\gamma}{\gamma-1}P_1 v_s + \frac{B_1^2}{8\pi}v_s = \frac{1}{2}\frac{\rho_1 v_s^3}{x^2} + \frac{\gamma}{\gamma-1}\frac{P_2 v_s}{x} + \frac{B_1^2}{8\pi}v_s x \tag{6.18}$$

One solution to these equations is the trivial one: $\rho_1 = \rho_2, u_1 = u_2, P_1 = P_2,$ and $B_1 = B_2$. But this is boring.

We can solve the modified momentum equation for P_2 and substitute into the modified energy equation to eventually get a quadratic in x. Draine lists the solution then as

$$x = \frac{2(\gamma+1)}{D + \sqrt{D^2 + 4(\gamma+1)(2-\gamma)M_A^{-2}}}\,, \tag{6.19}$$

where

$$D \equiv (\gamma-1) + 2M^{-2} + \gamma M_A^{-2} \tag{6.20}$$

$$M \equiv \frac{v_s}{\sqrt{\gamma P_1/\rho_1}} \tag{6.21}$$

$$M_A \equiv \frac{v_s}{B_1/\sqrt{4\pi\rho_1}}\,. \tag{6.22}$$

M is again the Mach number and M_A is the Alfven Mach number.

For a shock to exist, it must be supersonic, so $v_s > c_{ms}$ (Equation 6.3). We can then define yet another Mach number:

$$\mathcal{M} \equiv v_s/c_{ms}\,, \tag{6.23}$$

and this is the one that matters for a magnetized medium.

These equations don't reduce to a nice form unless we have a "strong shock:" $\mathcal{M} \gg 1$. In this case, $D \to (\gamma-1)$, so

$$x \to \frac{\gamma+1}{\gamma-1} = 4 \text{ for } \gamma = 5/3\,. \tag{6.24}$$

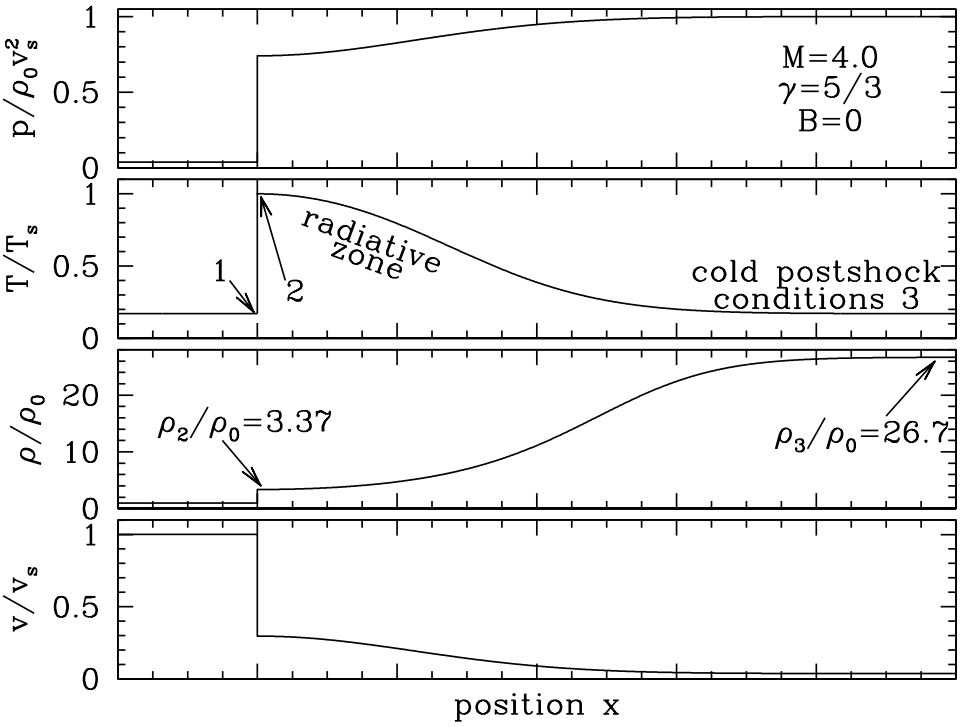


Figure 6.1: Draine figure 36.1 showing the structure of a nonmagnetic radiative shock with $M = 4$. Our treatment so far has only concerned positions 1 and 2.

It also follows that

$$u_2 \to \frac{\gamma - 1}{\gamma + 1} v_s = \frac{1}{4} v_s \text{ for } \gamma = 5/3\,. \tag{6.25}$$

If we then solve for the pressure P_2 and assume the ideal gas law $T_2 = P_2\mu/\rho_2 k$, then

$$T_2 \to \frac{2(\gamma - 1)}{(\gamma + 1)} \frac{\mu v_s^2}{k} = \frac{3}{16} \frac{\mu v_s^2}{k} \text{ for } \gamma = 5/3\,. \tag{6.26}$$

Draine provides handy values for T_2:

$$T_2 \approx 2890\,\mathrm{K} \left(\frac{\mu}{1.273 m_H} \right) \left(\frac{v_s}{10\ \mathrm{km\,s^{-1}}} \right)^2 \tag{6.27}$$

$$T_2 \approx 1.38 \times 10^7\,\mathrm{K} \left(\frac{\mu}{0.609 m_H} \right) \left(\frac{v_s}{1000\ \mathrm{km\,s^{-1}}} \right)^2 , \tag{6.28}$$

where $\mu = 1.273 m_H$ for H I and $\mu = 0.609 m_H$ for fully ionized gas.

We can see these effects graphically in Draine Figure 36.1, for an unmagnetized flow.

6.4 Magnetic fields

We hate magnetic fields, because they are difficult to measure and mess up all of our nice equations. Unfortunately, they can be very important in dynamic processes, as we saw in

the preceding discussion of shocks.

Magnetic fields pervade all size scales in astrophysics. They are especially important in dynamic processes, since they add pressure that (generally) resists gravitational contraction. In cgs, this pressure, or energy density, is

$$P_B = \frac{B^2}{8\pi} \tag{6.29}$$

Draine lists that for a median Galactic magnetic field strength of 6.0 μG (Heiles & Crutcher, 2005), the magnetic energy density is 0.89 eV cm^{-3}. Magnetic fields therefore have a higher energy density than almost anything else averaged over the entire Galaxy (Draine Table 1.5).

Magnetic fields essentially do not interact with neutral gas, although as we learned in the chapter on dust there can be small effects. Neutral gas that is pervaded by a strong magnetic field will act the same as if there is no strong magnetic field. Plasmas will feel an additional force along the field lines. They are most important for charged particles. They can, however, still have notiable effects on neutral gases with small ionized components. Luckily (or not!), most astrophysical systems have enough ionization to "freeze" the magnetic field to the plasma. This ionization is maintained by UV photons in the diffuse ISM, and by cosmic rays in denser clouds. If the plasma moves, it carries the magnetic field with it, which is called "advection."

An important quantity for magnetic fields is the "plasma beta," which tells you the ratio of turbulent to magnetic energy densities (pressures):

$$\beta = \frac{8\pi nkT}{B^2}\,. \tag{6.30}$$

In high-β environments, thermal energy dominates. In low-β environments, magnetic energy dominates.

https://arxiv.org/pdf/1001.5230v2.pdf

6.4.1 Observations of magnetic fields

Magnetic fields are notoriously difficult to measure, and furthermore such measurements never give you the total magnetic field, typically only the line of sight component. Despite being dynamically important, researchers often try to ignore magnetic fields because of our ignorance of their strenghts.

Magnetic field observations in most cases (all?) require observations of polarized light. If the observations are done in full Stokes, we get I, Q, U, V (hopefully you covered this in the radio astronomy course). We can get the direction of linearly polarized emission from $Q = I\cos(2\psi)$ and $U = I\sin(2\psi)$. The linearly polarized intensity is $PI = \sqrt{Q^2 + U^2}$.

[From Cal. Tech course http://www.its.caltech.edu/∼kamion/Ay126/Bfields.pdf

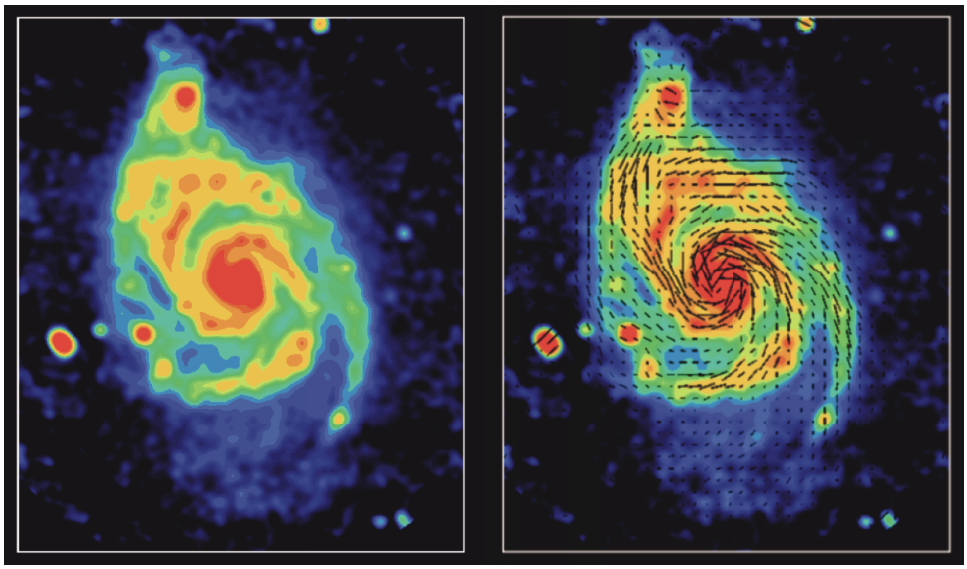


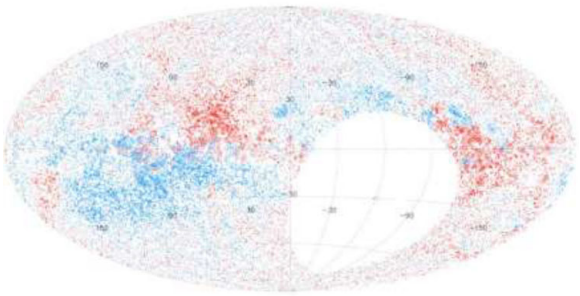
Figure 6.2: The total radio continuum emission from the "whirlpool" galaxy M51 (distance estimates range between 13 and 30 million light years) is strongest at the inner edges of the optical spiral arms, probably due to the compression of magnetic fields by density waves. The vectors give the orientations of the regular magnetic fields as derived from the polarized emission. The field lines follow nicely the optical spiral arms. Unexpectedly, strong polarized emission is observed also between the optical arms which indicates the action of a dynamo. This image was observed with the VLA in its most compact configuration at 6cm radio wavelength (broadband continuum). As the VLA cannot detect the diffuse, large-scale radio emission, data from the Effelsberg 100-m telescope in Germany at the same wavelength was added. Investigator(s): Rainer Beck (MPIfR Bonn, Germany), Cathy Horellou (Onsala Space Observatory). http://images.nrao.edu/336

Faraday rotation

As we saw previously, plasmas have a frequency-dependent index of refraction. The speeds of electromagnetic waves that propagate through that plasma thus have a frequency-dependent group velocity. If there is a source (e.g., a pulsar) that emits pulsed radiation, then the arrival times of those pulses will depend on the frequency (and the time delay depends on ν^{-2}, where ν is the frequency). Measurement of these time delays can be used to infer the dispersion measure.

If the plasma is magnetized, then there is an additional effect that acts on the polarization of the electromagnetic wave. The Faraday rotation effect appears during the propagation of electromagnetic waves in a magnetized plasma. A linearly polarized wave can be decomposed into opposite-handed circularly polarized components. The right-handed and left-handed circularly polarized waves propagate with different phase velocities within the magneto-ionic material. This effectively rotates the plane of polarization of the electromagnetic wave.

Suppose there is a magnetic field in the plasma directed along the line of sight. Electrons will then spiral in one particular direction around this magnetic field, and so the indexes



of refraction for right- and left-circularly polarized electromagnetic waves will differ. The propagation speeds for right- and left-circularly polarized waves will therefore differ slightly. Recalling that a linearly-polarized wave is a superposition of two circularly-polarized waves, the linearly-polarized wave will undergo something like a beat phenomenon that occurs when two waves of slightly different frequencies are superposed. What this results in is a rotation of the linear polarization of a linearly polarized EM wave by an angle, $\Psi = RM\lambda^2$, where λ is the wavelength and

$$RM = \frac{1}{2\pi}\frac{e^3}{m_e^2 c^4}\int_0^L n_e B_\parallel dL \tag{6.31}$$

If the DM and RM are both measured, after evaluating the constants the electron-density-weighted mean line-of-sight magnetic field is

$$\langle B_\parallel \rangle = \frac{RM}{8.12\times10^{-5}\mathrm{rad\,cm^{-2}}}\frac{\mathrm{cm^{-3}\ pc}}{DM}\,\mu\mathrm{G}\,. \tag{6.32}$$

This can be measured along many different lines of sight, and also to pulsars at different distances along similar lines of sight, to get information about the three-dimensional magnetic field. Measurements indicate magnetic fields $B \sim 2-4\,\mu$G in the spiral arms and slightly smaller in the interarm regions, with a sign flip between arm and interarm (note that $6\,\mu$G listed previously).

We usually measure differences in Ψ to determine the rotation measure:

$$RM = \frac{\Psi_2 - \Psi_1}{\lambda_1^2 - \lambda_1^2} \tag{6.33}$$

This shows the π-ambiguity from two wavelengths alone. With more wavelengths, we can resolve this ambiguity.

Synchrotron radiation

We learned before that the strength of synchrotron radiation depends on the magnetic field strength. We had that the source function is:

$$S_\nu = \frac{j_\nu}{\kappa_\nu} \propto B^{-1/2}\nu^{5/2}\,. \tag{6.34}$$

At low frequencies, the synchrotron radiation is optically thick, and $I_\nu \propto S_\nu \propto B^{-1/2}\nu^{5/2}$. At high frequencies, the intensity will be proportional to the emissivity coefficient: $I_\nu \propto j_\nu \propto \nu^{-(\gamma-1)/2} \propto \nu^{-0.7}$. Note that I have switched our electron spectral index symbol from the p used earlier to γ, because we need p to define polarization below.

Synchotron radiation is also polarized if the magnetic fields are coherent. If the magnetic field is homogeneous, then the polarization is $p = (\gamma + 1)/(\gamma + 7/3) \simeq 0.73$. Observations show $p = 0.1 - 0.2$, typically, a consequence of incoherence or Faraday depolarization, beam effects, etc. Still, the variation of the polarization across the sky can be used to infer a magnetic-field pattern in the Galaxy.

Zeeman splitting

A magnetic field splits the electronic energy states of the same l into $2l + 1$ energy levels split by energies $\sim \mu_B B \simeq 5.78 \times 10^{-15}\ (B/\mu\mathrm{G})$ eV, where $\mu B = e\hbar/2m_e c$ is the Bohr magneton. Magnetic fields of strength $1 - 100\,\mu$G give rise to level splittings that are too small to be detected in sub-mm or shortward ($h\nu \gtrsim 10^{-4}\,$eV). The splitting in the 21-cm line ($h\nu = 5.9 \times 10^{-6}\,$eV) is split by $\sim 10^{-8}$ for a $10\,\mu$G field. This is smaller than the $v/c \simeq 10^{-5}$ frequency shifts from velocity broadening in molecular clouds or the IGM and thus unobservable.

However, Zeeman splitting gives rise to a difference in the frequencies of the two circular polarizations of the transition radiation, and these can be detected and have been used to measure magnetic-field strengths in H I and molecular regions in the ISM. Diffuse clouds studied in 21-cm absorption this way have been found to have $B \simeq 6\,\mu$G implying a magnetic pressure several times larger than the gas pressure. Thus, magnetic fields may be dynamically important in H I regions.

After Dick Crutcher, Zeeman diagrams are callsed "Crutcher diagrams." They show that in low density ionized gas (including partially ionized H I), there is no relationship between magnetic field strength and the density. For molecular clouds, there is a linear relationship. This shows that as clouds collapse, the field is advected. Therefore, the field strength and density increase together. The Zeeman effect is line of sight B-field, so the points generally fal under the curves.

Polarization of starlight

Evidence that polarization of starlight was due to interstellar dust came from (a) the correlation between the magnitude of the polarization and the reddening, and (b) the coherence in the polarization between different stars in the same region of the sky. We know from before that in most cases the polarization percentage peaks near the B band (5500 Å) and follows the empirical Serkowski law,

$$p(\lambda) \approx p_{\max} e^{-K \ln^2(\lambda/\lambda_{\max})}\,, \tag{6.35}$$

with $\lambda_{\max} \approx 500\,$nm and $K \approx 1.15$.

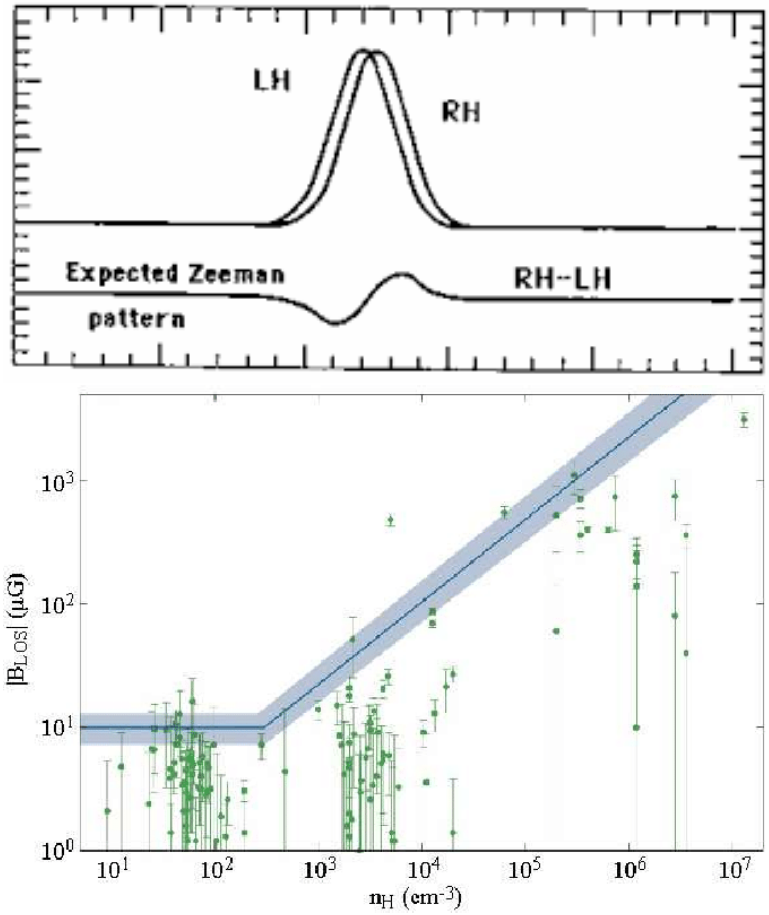


Figure 6.3: Zeemann profile (top) and Crutcher diagram (bottom).

Heuristically, the polarization arises if dust grains are elongated and somehow aligned. If so, then the absorption of light polarized along the long axis may differ from that along the short axis. The dust grains appear to be aligned by the interstellar magnetic field with their shortest axes parallel to the magnetic field; the mechanism is not well understood and is a subject of active current research. Since extinction increases toward the UV, while the polarization decreases, it suggests that the grains responsible for the polarization have radii $a \simeq 2(\lambda_{\rm max}/2\pi) \simeq 0.1\,\mu$m. In the UV, one moves to the geometric-optics limit, and both polarizations are absorbed similarly. Thus, the V band extinction must be due largely to $a \simeq 0.1\,\mu$m grains, and these grains must be nonspherical and aligned with the magnetic field; and grains with $a \lesssim 0.05\,\mu$m, which dominate the exinction at $\lambda \lesssim 0.3\,\mu$m are either spherical (which seems unlikely) or minimally aligned.

6.5 Supernovae and the Three-Phase ISM

Supernovae (SNe) are super explosions of super stars at the end of their lifetimes. They are one of the most energetic phenomenae in the Universe, and have a large impact on driving turbulence in the ISM. They are believed to be responsible for the acceleration of Galactic cosmic rays and the creation of the HIM.

As a review: SN Type Ia are from white dwarf accretion past the Chadrasehkar limit of $\sim 1.4\ M_\odot$. All other SN are from core-collapse, including the famous SN 1987A in the LMC.

Video of SN1987A:
https://www.youtube.com/watch?v=xigKhIfD_Ko
https://en.wikipedia.org/wiki/SN_1987A#/media/File:SN1987a_debris_evolution_animation_time_scaled.g

A typical SN has an energy of $E_0 = 10^{51}\,$erg ($E_{51} = 1$), although some Type II SNe have $E_0 = 10^{52}\,$erg. The ejected mass ($M_{\rm ej}$) ranges from $\sim 1.4\ M_\odot$ for Type I SNe, and $\sim 10 - 20\ M_\odot$ for Type II SNe.

We can understand SNe in terms of some very simple physics, and break things into three distinct phases: free expansion,

6.5.1 Phase I: Free Expansion

The first phase of expansion, the expansion energy is significantly greater than the energy contained in the local medium. For type II supernovae, the expansion initially moves into the exterior parts of the star.

Let's compare the ejecta energy with that of the explosion itself to see how fast this phase is propagating into the ISM:

$$\left\langle v_{ej}^2 \right\rangle = \left(\frac{2E_0}{M_{\rm ej}}\right)^{1/2} = 1.00 \times 10^4\ {\rm km\,s^{-1}}\ E_{51}^{1/2} \left(\frac{M_\odot}{M_{\rm ej}}\right)^{1/2} . \tag{6.36}$$

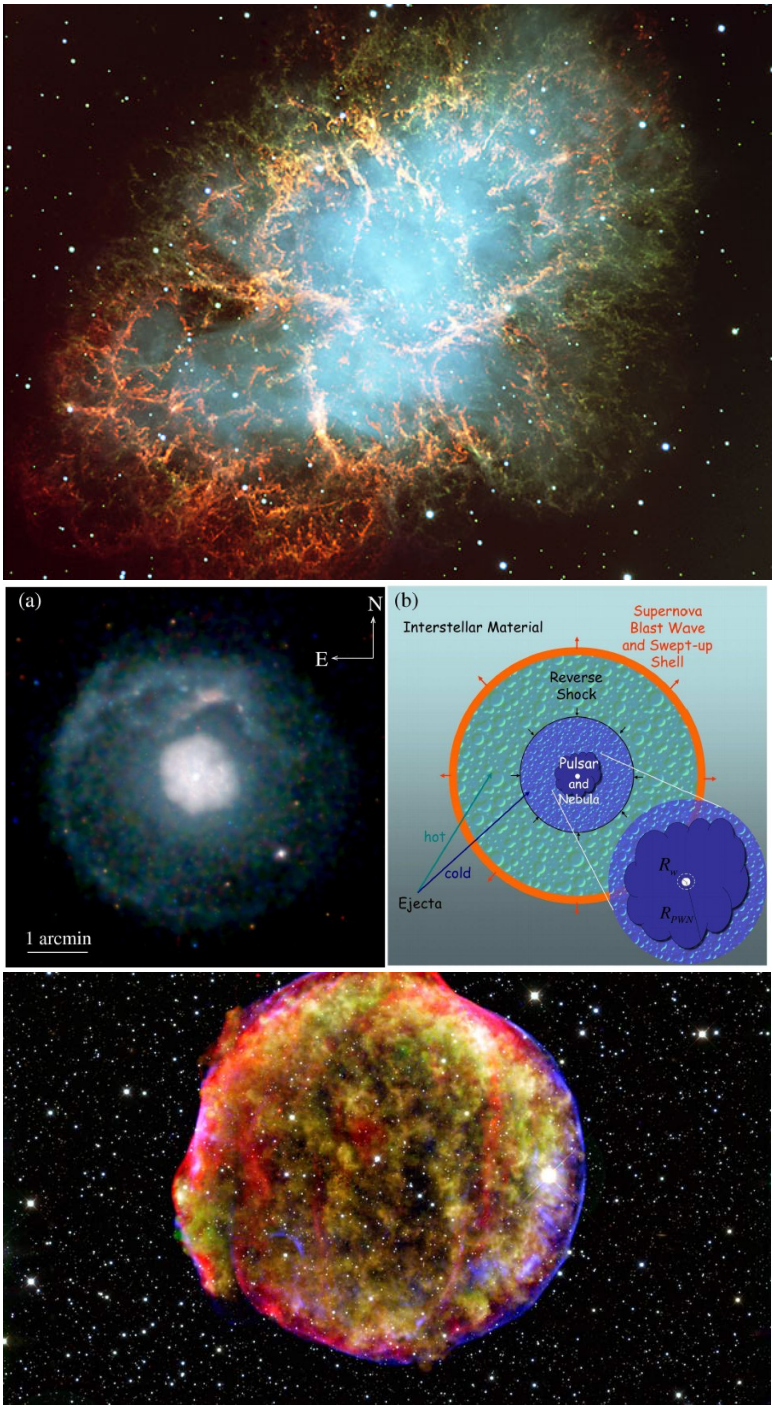


Figure 6.4: Crab nebula in the optical, diagram of blast wave and reverse shock, Tycho's SNR at X-ray wavelengths.

This is obviously much greater than the local sound speed of a few $\mathrm{km\,s^{-1}}$, which leads to a fast shock expanding into the ISM. Interior to the shock is the supernova remnant (SNR). As long as the material swept up by the shock is much less than the mass of the stellar ejecta, the expansion of the stellar ejecta proceeds at essentially a constant velocity equal to

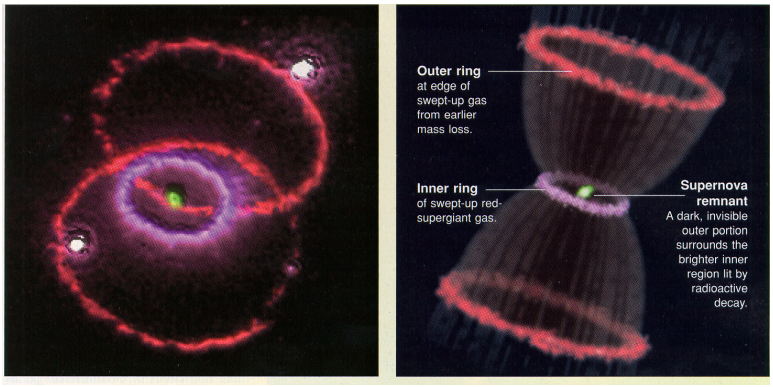


Figure 6.5: SN1987A from *Sky & Telescope*

the initial shock wave speed, typically of the order of 10,000 $\mathrm{km\,s^{-1}}$. This is known as the "free expansion" phase and may last for approximately 200 years, at which point the shock wave has swept up as much interstellar material as the initial stellar ejecta. The supernova remnant at this time will be about 3 pc in radius.

Although the remnant is radiating thermal X-ray and synchrotron radiation across a broad range of the electromagnetic spectrum (from radio to X-rays), the initial energy of the shock wave will have diminished very little. Line emission from the radioactive isotopes generated in the supernovae contribute significantly to the total apparent brightness of the remnant in the early years, but do not significantly affect the shock wave.

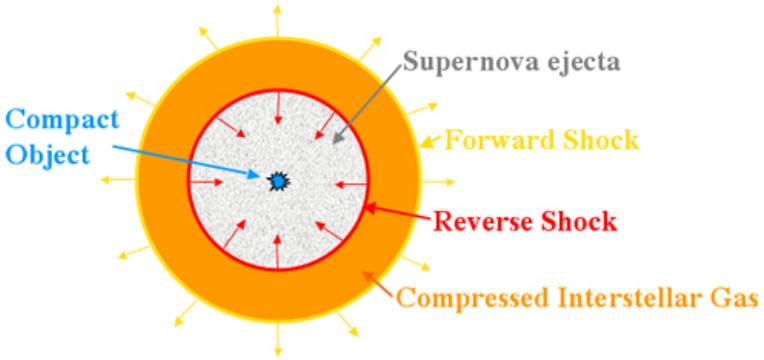
6.5.2 Phase II: Sedov-Taylor: The Blastwave$^{\mathrm{TM}}$

As the remnant sweeps up ambient mass equal to the mass of the stellar ejecta, the wave will begin to slow and the remnant enters a phase known as adiabatic expansion, or the Sedov-Taylor or blast wave phase. The internal energy of the shock continues to be very large compared to radiation losses from thermal and synchrotron radiation, so the total energy remains nearly constant. The rate of expansion is determined solely the initial energy of the shock wave and the density of the interstellar medium.

As the density of the expanding ejecta drops (as T^{-3}), the pressure of the shocked gas behind the shock wave soon exceeds the thermal pressure in the ejecta. Because of this pressure difference, a reverse shock is created. There are now two shock fronts. The original one propagating outward is called the "blastwave" and the reverse shock propagating inward. The reverse shock re-heats the material in the SNR. [For the interested reader: Can we compute when the reverse shock is created?]

As the ejecta expands out from the star, it passes through the surrounding interstellar medium, heating it from 10^7 to 10^8 K, sufficient to separate electrons from their atoms and to generate thermal X-rays. The interstellar material is accelerated by the shock wave and will be propelled away from the supernova site at somewhat less than the shock wave's initial velocity. This makes for a thin expanding shell around the supernova site encasing a

relatively low density interior.



This occurs at radius:

$$R_1 = \left(\frac{3M_{\rm ej}}{4\pi\rho_0}\right)^{1/3}, \tag{6.37}$$

when

$$t_1 \approx \frac{R_1}{\langle v_{\rm ej}^2\rangle^{1/2}} = 186\,{\rm yr}\left(\frac{M_{\rm ej}}{M_\odot}\right)^{5/6} E_{51}^{-1/2} n_0^{-1/3} \tag{6.38}$$

For $t \geq t_1$, the reverse shock has already reached the center of the SNR, and the entire SNR is hot. The remnant is still expanding due to the large pressure difference between the ISM and the SNR. It is emitting, but this radiation is not cooling the remnant significantly because the densities are low.

We can now approximate the expansion as a "point explosion" injecting energy E_0 into the uniform density ISM. We can neglect the finite mass of the ejecta (which is dwarfed by the mass of the swept up material), the radiative losses of energy (which are small compared to the energy of the system), and the pressure in the ambient medium (small compared to that in the SNR).

Here Draine switches to a strange dimensional analysis method to arrive at a classic result. We know that the shock radius R_s will expand at a rate dependent on the SN energy and the mass of the ISM:

$$R_s = AE^\alpha \rho^\beta t^\eta\,, \tag{6.39}$$

where the explosion occurs at $t = 0$. From dimensional analysis:

$$\begin{aligned} {\rm Mass}: 0 &= \alpha + \beta & (6.40)\\ {\rm Length}: 1 &= 2\alpha + 3\beta & (6.41)\\ {\rm Time}: 0 &= -2\alpha + \eta\,. & (6.42) \end{aligned}$$

The mass condition arises because energy and mass must be proportional. To get length out, their exponents must differ only by a sign. The length condition arises because energy has in its units length^{-2} and density as length^{-3}; if $\alpha = -1$, $\beta = 1$ to get length out of the

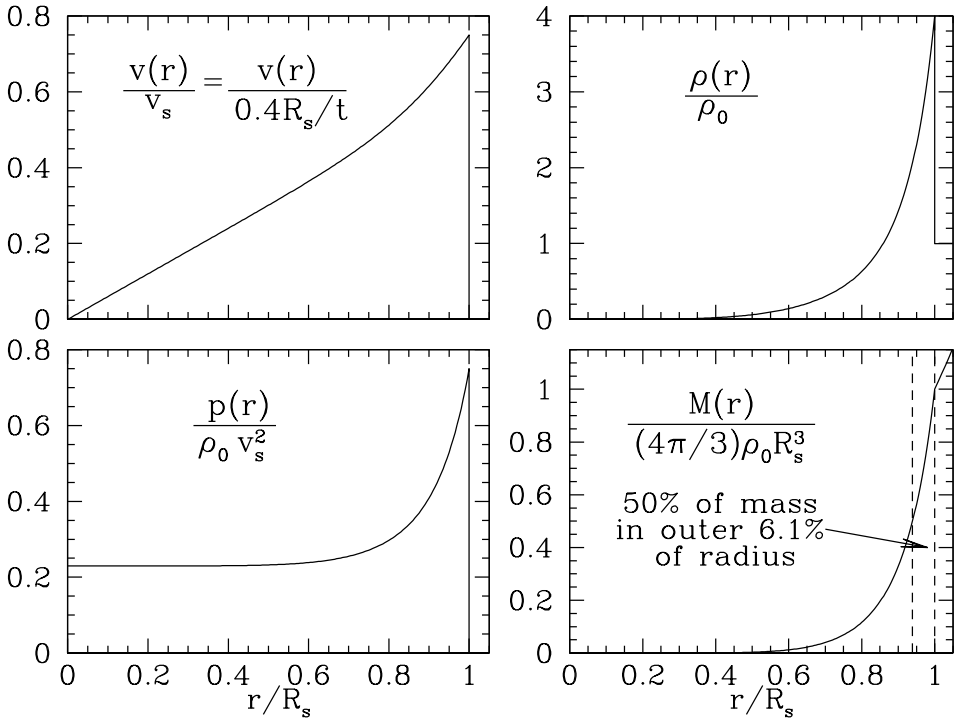


Figure 6.6: Sedov-Taylor expansion for $\gamma = 5/3$ [Draine 39.1].

"equation." Similarly, for the third condition, energy has in its units time^{-2}, so if $\alpha = 1$, $\eta = 2$. These are not the final values though, just the proportions.

This leaves us with three equations and three unknowns. We can easily solve these to get $\alpha = 1/5, \beta = -1/5, \eta = 2/5$:

$$R_s = A \left(\frac{Et^2}{\rho_0} \right)^{1/5} , \tag{6.43}$$

where $A = 1.15$ from the exact solution. Neat! We can therefore rewrite our expansion terms, after realizing that $v^2 \propto T$:

$$R_s = 1.52 \times 10^{19} \text{ cm } E_{51}^{1/5} n_0^{-1/5} t_3^{2/5} \tag{6.44}$$

$$v_s = 1950 \text{ km s}^{-1} E_{51}^{1/5} n_0^{-1/5} t_3^{-3/5} \tag{6.45}$$

$$T_s = 5.25 \times 10^7 \text{ K } E_{51}^{2/5} n_0^{-2/5} t_3^{-6/5} , \tag{6.46}$$

Or, the radius grows slowly with time, the shock velocity decreases slowly with time, and the temperature decreases with time.

Draine shows the formal Sedov-Tailer solution.

Draine mentions that the Sedov-Taylor solution is not too bad, although it does neglect some dynamical effects.

6.5.3 Phase III: Snowplow phase: Escape from Sedov-Taylor: The Reckoning: The Radiative Phase

[When does the Sedov-Taylor expansion phase end?] When radiative cooling becomes important. When temperatures cool to about $20,000$ K, ions and electrons begin recombining, the SNR leaves the Sedov-Taylor expansion phase.

This is probably a good time to talk about the radiation. The SNR is $\sim 10^7$ K. How does it radiate? X-rays and synchrotron primarily. Why not free-free? Well, there is free-free as well, but the synchrotron emission is much stronger.

After the temperature cools, the hot recombined electrons emit UV line radiation. This is much more efficient at cooling the remnant. For a cooling function Λ that has units of erg s^{-1} cm^{-3}, this gives

$$\frac{dE}{dt} = -\int_0^{R_s} \Lambda \, 4\pi r^2 \, dr \tag{6.47}$$

Draine mentions a functional form for Λ that leads to:

$$t_{\rm rad} = 49.3 \times 10^3 \, {\rm yr} \, E_{51}^{0.22} n_0^{-0.55} \tag{6.48}$$

$$R_{\rm rad} = 7.32 \times 10^{19} \, {\rm cm} E_{51}^{0.29} \tag{6.49}$$

When $t \approx t_{\rm rad}$, the thermal pressure behind the shock has dropped significantly due to cooling. We call this the snowplow phase. There is now a dense shell of cool gas that is enclosing a hot central volume. The snowplow here refers to the fact that the dense shell mass is added to as the blastwave progresses outward.

The gas here is just cooling by adiabatic expansion. Adiabatic here just means that the gas does not transfer heat to its surroundings (via radiation). Hence this new phase is known as the radiative phase during which X-ray radiation becomes much less apparent and the remnant cools and disperses into the surrounding medium over the course of the next 10000 years.

At the beginning of the snowplow phase, Draine notes that the shock speed is ~ 150 km s^{-1}.

Draine gives the relevant expressions for the snowplow phase:

$$R_s \approx R_s(t_{\rm rad}) \left(\frac{t}{t_{\rm rad}} \right)^{2/7} \tag{6.50}$$

$$v_s \approx \frac{2}{7} \frac{R_s}{t_{\rm rad}} \left(\frac{t}{t_{\rm rad}} \right)^{-5/7} \tag{6.51}$$

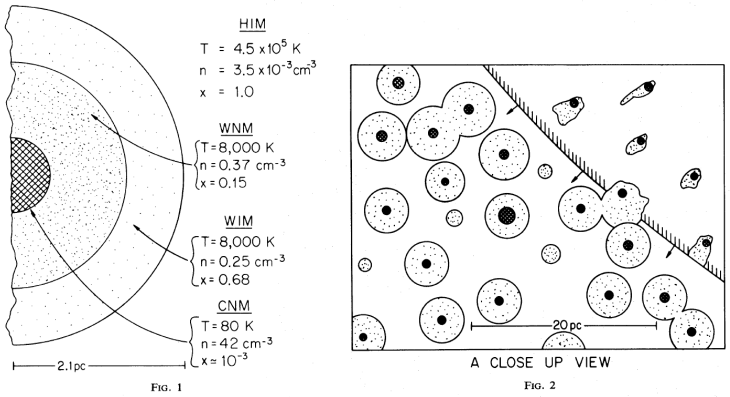


FIG. 1 FIG. 2

6.5.4 Phase IV: Fadeaway

The shock speed declines with time until it becomes just an ordinary sound wave. Using our previous expressions, this occurs when

$$t_{\rm fade} \approx \left(\frac{(2/7)R_{\rm rad}/t_{\rm rad}}{c_s}\right)^{7/5} t_{\rm rad} \approx 1.87\times 10^6\,{\rm yr}\; E_{51}^{0.32} n_0^{-0.37} \left(\frac{c_s}{10\;{\rm km\,s^{-1}}}\right)^{-7/5} \tag{6.52}$$

$$R_{\rm fade} \approx 2.07\times 10^{20}\;{\rm cm}\, E_{51}^{0.32} n_0^{-0.37} \left(\frac{c_s}{10\;{\rm km\,s^{-1}}}\right)^{-2/5} \tag{6.53}$$

Why does this happen? Internal pressure is not greater than external pressure. The SN could run into a dense structure, or radiative cooling may dominate.

6.5.5 Why would we care about this?

McKee & Ostriker (1977), in a classic paper, argued that blastwaves from SNe have a large impact on shaping the ISM. The envisioned an ISM consisting primarily of the CNM and the HIM. The WNM and WIM are restricted to the interface regions of the neutral clouds, and the WIM in direct contact with the HIM and photoionized by thermal emission from it. A blastwave propagates into these media.

The authors view the ISM as being composed of numerous small (spherical!) clouds of molecular gas, embedded in a diffuse hot ISM (HIM). Each cloud has an ionized halo (the WIM) maintained by the interstellar UV background. Between the ionized halo and the cloud itself, they suggest the presence of a neutral zone heated by interstellar X-rays.

It turns out this isn't really correct in detail, but nonetheless provides a useful framework.

6.6 Star Formation [Draine Chapter 41]

Stars have to form somehow, because stars exist. The formation of stars must involve collapse of a molecular cloud. These clouds have mean densities of $\sim 10^3$ cm^{-3} or so, and sizes of about a parsec ($\sim 10^{18}$ cm). They must collapse down to $\sim 10^{11}$ cm (the size of a star), with densities of $\sim 10^{24}$ cm^{-3} (the mean density of a star).

When does this collapse occur? When gravity overcomes pressure. The condition where gravity and pressure are in balance is of course called "hydrostatic equilibrium." One treatment says that a cloud not in hydrostatic equilibrium that will collapse has a characteristic size of the "Jeans radius" and mass of the "Jeans Mass" (the condition of instability is the "Jeans Instability"). We will derive these quantities first from the hydrostatic equilibrium condition. We can also use the Virial theorem, and the equations of fluid mechanics (in Draine). All three derivations are important, and we'll do them all in turn.

6.6.1 Jeans Mass from Hydrostatic Equilibrium [Following Wikipedia page]

The Jeans mass is named after the British physicist Sir James Jeans, who considered the process of gravitational collapse within a gaseous cloud. He was able to show that, under appropriate conditions, a cloud, or part of one, would become unstable and begin to collapse when it lacked sufficient gaseous pressure support to balance the force of gravity. The cloud is stable for sufficiently small mass (at a given temperature and radius), but once this critical mass is exceeded, it will begin a process of runaway contraction until some other force can impede the collapse. He derived a formula for calculating this critical mass as a function of its density and temperature. The greater the mass of the cloud, the smaller its size, and the colder its temperature, the less stable it will be against gravitational collapse.

Hydrostatic equilibrium is:

$$\frac{dp}{dr} = -\frac{G\rho(r)M_{enc}(r)}{r^2}, \tag{6.54}$$

where $M_{enc}(r)$ is the enclosed mass, p is the pressure, $\rho(r)$ is the density of the gas at r, G is the gravitational constant and r is the radius. The equilibrium is stable if small perturbations are damped and unstable if they are amplified. In general, the cloud is unstable if it is either very massive at a given temperature or very cool at a given mass for gravity to overcome the gas pressure.

Let's say we have a spherical molecular cloud cloud of radius R, mass M, and sound speed c_s. Compression of this region can only proceed at approximately the sound speed, which gives a characteristic time of:

$$t_{\rm sound} = \frac{R}{c_s} \tag{6.55}$$

for sound waves to cross the region. Gravity will attempt to contract the system even further, and will do so on a free-fall time,

$$t_{\rm ff} = \sqrt{\frac{3\pi}{32G\rho}}\,. \tag{6.56}$$

(This is sometimes given as $t_{\rm ff} = \sqrt{\frac{3}{2\pi G\rho}}$, from a simpler treatment.) It is important to remember that $t \approx (G\rho)^{-1/2}$ is the characteristic time for many processes in astrophysics. This is a good starting guess for many time scales.) We have collapse when $t_{\rm ff} < t_{\rm sound}$. In this case, the collapse is fast enough that the cloud cannot re-establish equilibrium, which takes place over the timescale given by the sound speed.

It is worth here taking a slight detour to describe how long these free fall times are. For large scales, the growth time for the Jeans instability is

$$\tau_J \simeq 2.3 \times 10^4 \mathrm{yr} \left(\frac{10^6 \text{ cm}^{-3}}{n_H} \right)^{1/2} \tag{6.57}$$

For $n_H = 1000$ cm^{-3}, this is about 0.7 Myr. Free fall time (collapse timescale for a pressureless gas) is:

$$\tau_{ff} = \left(\frac{3\pi}{32 G \rho_0} \right)^{1/2} = 4.4 \times 10^4 \,\mathrm{yr} \left(\frac{10^6 \text{ cm}^{-3}}{n_H} \right)^{1/2} \tag{6.58}$$

For $n_H = 1000$ cm^{-3} this is 1.4 Myr - slightly longer than growth time.

OK, back to the Jeans mass and radius. The resultant Jeans radius R_J is therefore:

$$\lambda_J \simeq \frac{c_s}{\sqrt{G\rho}} \tag{6.59}$$

The speed of sound is

$$c_s = \sqrt{\frac{\gamma P}{\rho}}\,, \tag{6.60}$$

where γ is the adiabatic index, which is 7/5 for molecular gas and 5/3 for monotonic gas. The pressure $P = nkT = \rho/\mu kT$ assuming an ideal gas, with mean molecular mass μ, so we have

$$R_J \simeq \sqrt{\frac{kT}{G\mu\rho}}\,. \tag{6.61}$$

The real definition gives a factor of order unity out front:

$$R_J \simeq \sqrt{\frac{15kT}{4\pi G\mu\rho}} \simeq (0.4 \text{ pc}) \left(\frac{c_s}{0.2 \text{ km s}^{-1}} \right) \left(\frac{n}{10^3 \text{ cm}^{-3}} \right)^{-1/2}\,. \tag{6.62}$$

All scales larger than the Jeans length are unstable to gravitational collapse, whereas smaller scales are stable.

Perhaps the easiest way to conceptualize Jeans Length is in terms of a close approximation, in which we rephrase ρ as M/r^3. The formula for Jeans' Length then becomes:

$$R_J \approx \sqrt{\frac{k_B T r^3}{GM\mu}}\,, \tag{6.63}$$

and therefore $R_J = r$ when $kT = \frac{GM\mu}{r}$. In other words, the cloud's radius is the Jeans Length when thermal energy per particle equals gravitational work per particle. At this critical length the cloud neither expands nor contracts. It is only when thermal energy is not equal to gravitational work that the cloud either expands and cools or contracts and warms, a process that continues until equilibrium is reached.

We can recast this in terms of the "Jeans mass":

$$M_J = \left(\frac{4\pi}{3}\right)\rho R_J^3 = \left(\frac{5kT}{G\mu}\right)^{3/2}\left(\frac{3}{4\pi\rho}\right)^{1/2} \simeq (2\ \mathrm{M}_\odot)\left(\frac{c_s}{0.2\ \mathrm{km\ s^{-1}}}\right)^3\left(\frac{n}{10^3\ \mathrm{cm^{-3}}}\right)^{-1/2}. \tag{6.64}$$

The Jeans mass M_J is just the mass contained in a sphere of radius R_J. It is useful to remember that $M_J \propto T^{3/2}\rho^{-1/2}$. Thus, stars can form most efficiently (when mass is low) in low temperature, high density locations where the Jeans mass is not as great.

The above is an illustrative and wrong derivation! Jeans assumed that the collapsing region of the cloud was surrounded by an infinite, static medium. The pressure in hydrostatic equilibrium is therefore less than that required, and the mass is therefore too high. We will fix this problem below.

A larger issue is that because all scales greater than the Jeans length are also unstable to collapse, any initially static medium surrounding a collapsing region will in fact also be collapsing. As a result, the growth rate of the gravitational instability relative to the density of the collapsing background is slower than that predicted by Jeans' original analysis. This flaw has come to be known as the "Jeans swindle".

6.6.2 The Jeans Mass from the Virial Theorem

We can also derive the Jeans mass using the Virial theorem. Like the condition of hydrostatic equilibrium, the Virial theorem describes a system in equilibrium. If the kinetic energy of a system is K and the gravitational potential energy is U, the simplest incarnation of the Virial theorem says that $2K + U = 0$.

Each particle in gas cloud has kinetic energy, $E = 3/2kT$, so the total kinetic energy $K = \Sigma_i^N E_i = 3/2NkT$, where N is the number of particles.

For an isothermal sphere,

$$U = -\frac{3}{5}\frac{GM^2}{R}, \tag{6.65}$$

so

$$3NkT = -\frac{3}{5}\frac{GM^2}{R}. \tag{6.66}$$

We can replace $N = M/\mu$, with μ the mass per particle, and $R = (3M/4\pi\rho)^{1/3}$ to get

$$M_J = \left(\frac{5kT}{G\mu}\right)^{3/2}\left(\frac{3}{4\pi\rho}\right)^{1/2} \tag{6.67}$$

The same as before!

6.6.3 The Jeans Mass from the Fluid Equations [Draine Chapter 41, but following Harvard ISM notes here]

Start with the basic hydro equations (conservation of mass and momentum, plus Poisson's equation for the gravitational potential):

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \vec{v}) = 0 \text{ [Conservation of mass]} \quad (6.68)$$

$$\frac{\partial v}{\partial t} + (\vec{v} \cdot \nabla)\vec{v} = -\frac{1}{\rho}\nabla P - \nabla \Phi \text{ [Conservation of momentum]} \quad (6.69)$$

$$\nabla^2 \Phi = 4\pi G \rho \text{ [Poisson's Equation for the Gravitational Potential]}. \quad (6.70)$$

Φ is the gravitational potential.

Consider an equilibrium solution $\rho_0(\vec{r}), P_0(\vec{r})$, etc such that time derivatives are zero. Let's perturb that solution slightly, and analyze when that perturbation grows unstably: $\vec{v} = \vec{v_0} + \vec{v_1}, \rho = \rho_0 + \rho_1, P = P_0 + P_1, \Phi = \Phi_0 + \Phi_1$.

The linear hydro equations, to first order in the perturbations (getting rid of products of perturbed quantities), are

$$\frac{\partial \rho_1}{\partial t} + \vec{v_0} \cdot \nabla \rho_1 + \vec{v_1} \cdot \nabla \rho_0 = -\rho_1 \nabla \cdot \vec{v_0} - \rho_0 \nabla \cdot \vec{v_1} \text{ [Perturbed conservation of mass]} \quad (6.71)$$

$$\frac{\partial v_1}{\partial t} + (\vec{v_0} \cdot \nabla)\vec{v_1} + (\vec{v_1} \cdot \nabla)\vec{v_0} = \frac{\rho_1}{\rho_0^2}\nabla P_0 - \frac{1}{\rho_0}\nabla P_1 - \nabla \Phi_1 \text{ [Perturbed conservation of momentum]} \quad (6.72)$$

$$\nabla^2 \Phi_1 = 4\pi G \rho_1 \text{ [Perturbed Poisson's Equation]}. \quad (6.73)$$

For an isothermal gas, the equation of state is $P = \rho c_s^2$, where c_s is the isothermal sound speed. Then, the momentum becomes

$$\frac{\partial \vec{v_1}}{\partial t} + (\vec{v_0} \cdot \nabla)\vec{v_1} + (\vec{v_1} \cdot \nabla)\vec{v_0} = -c_s^2 \nabla (\frac{\rho_1}{\rho_0}) - \nabla \Phi_1 \quad (6.74)$$

Jeans took these equations and added:
Uniform density to start with ($\nabla \rho_0 = 0$)
Stationary gas ($v_0 = 0$)
Gradient-free equilibrium potential $\nabla \Phi_0 = 0$)
Then, the solution becomes (after taking divergence of above equation, the momentum equation)

$$\frac{\partial^2 \rho_1}{\partial t^2} = c_s^2 \nabla^2 \rho_1 + (4\pi G \rho_0)\rho_1 \quad (6.75)$$

Now consider plane wave perturbations

$$\rho_1 \propto \exp(i(\vec{k} \cdot \vec{r}) - \omega t) \quad (6.76)$$

$$\omega^2 = k^2 c_s^2 - 4\pi G \rho_0 \quad (6.77)$$

Define $k_J^2 = 4\pi G \rho_0 / c_s^2$, so $\omega^2 = (k^2 - k_J^2)c_s^2$. ω is real IFF $k > k_J$. Otherwise, ω is imaginary, and there is exponential growth of the instability. This then leads to a Jeans Length:

$$\lambda_J = 2\pi/k_J = \left(\frac{\pi c_s^2}{G\rho_0}\right)^{1/2} \tag{6.78}$$

This is exactly what we had before! Converting the Jeans length into a radius (assuming a sphere) yields

$$M_J = 0.32 M_\odot \left(\frac{T}{10K}\right)^{3/2} \left(\frac{m_H}{\mu}\right)^{3/2} \left(\frac{10^6\ \mathrm{cm}^{-3}}{n_H}\right)^{1/2} \tag{6.79}$$

Let's plug in values for a dense core: $T = 10$ K, $\mu = 2.33$ amu, $n_H = 2 \times 10^5$ cm^{-3}. This yields $M_J = 0.2\ M_\odot$. If we instead plug in numbers appropriate for the mean conditions in a GMC, $T = 50$ K, $\mu = 2.33$ amu, $n_{\rm H} = 200$ cm^{-3}, we get $M_J = 70\ M_\odot$.

Note that, once gravitational collapse and heating set in, our isothermal sphere assumptions are no longer valid.

6.6.4 Fragmentation

Of course this is a simplification – a single cloud does not collapse down to r=0. What happens to complicate the collapse? As the cloud collapse, density rises. Since the collapse is isothermal, a rising density means the Jeans mass of the cloud is falling, so small pieces of the cloud start to collapse on their own. A rising density also means a declining free fall time, so these small dense clumps collapse faster than the overall cloud.

Instead of one giant cloud undergoing a monolithic collapse, the cloud fragments into small collapsing pieces. So what stops this fragmentation? As the density rises, the opacity rises. At some point during the collapse and fragmentation process, the opacity rises high enough that the energy created during the collapse is absorbed within the star itself – it begins to heat up. Since the energy is not lost from the cloud, we call this an adiabatic collapse. Higher temperature means higher pressures (the ideal gas law), which halt the free collapse of the star. Since the cloud absorbs all the gravitational energy of collapse, it heats up, and it starts to act like a blackbody.

At what mass does this happen? We can balance the rate of energy loss through gravitational collapse to the rate at which the cloud radiates blackbody energy, and, solving for the mass, we find $M \approx M_\odot$. In other words, collapse halts when the fragment masses reach star-like masses.

6.6.5 Bonner-Ebert Spheres [Draine 41.3.1]

In a more realistic scenario, the density is centrally peaked. In this case, the gravitational energy is

$$U = -\frac{3}{5} a \frac{GM^2}{R}\,, \tag{6.80}$$

where $a > 1$ for centrally peaked density profiles. Draine mentions Mouschovias & Spitzer (1976) find $a \approx 1.67$ for numerical models of clouds on the verge of collapse.

In our above consideration of the Virial theorem, we neglected external pressure and magnetic energy. If we consider the former, with the above modification to the gravitational potential, we arrive at the "Bonner-Ebert mass" (Bonner 1956; Ebert 1957):

$$M_{\rm BE}(p_0) = \frac{225}{32\sqrt{5}\pi}\frac{c_s^4}{(aG)^{3/2}}\frac{1}{\sqrt{p}_0} = 0.26\left(\frac{T}{10\ {\rm K}}\right)^2\left(\frac{10^6\ {\rm cm}^{-3}\ {\rm K}}{p_0/k}\right)^{1/2} M_\odot \tag{6.81}$$

Remember how we said that the Jeans mass neglected some rather important things? Well, the Bonner-Ebert mass is basically the same as the Jeans mass, $M_{\rm BE} \approx 1.18 M_J$. The 18% change is due to the fact that the cloud itself affects the hydrostatic equilibrium assumption before.

Given the typical temperature and pressures of molecular clouds, the Bonner-Ebert mass is about a Solar mass, so it is probably no surprise that this is the peak of the IMF.

Of course we still neglect the magnetic fields. The magnetic energies are similar to the kinetic energies, and so can contribute to the pressure. See Draine for a complete treatment.

6.6.6 Formation of Actual Stars

During the formation of stars, cores more or less free-fall collapse. The free fall time depends inversely on the density, so the central part collapses first, then the outer parts. What would provide resistance? [Pressure of course!] What would pressure be unimportant? [Cooling from molecular lines!]. This "cooling" of course releases energy that we can detect. This energy peaks in the sub-millimeter to far-infrared.

Class 0, I, II, and III Protostars [Draine 42.2]

We classify protostars based on their spectral index. We used spectral indices in the radio before to distinguish between thermal (free-free) and non-thermal (synchrotron) emission. Here, the spectral index is computed in infrared bands, and is used to distinguish young from old protostars.

The earliest protostars will be cold. Their spectral energy distributions (SEDs) will therefore peak at longer wavelengths. We see emission from the disk surrounding the protostar, and also from the protostar itself. Whether we see emission from the disk or the protostar depends on the disk inclination angle and the evolutionary state - since the disk will dissipate with time.

The spectral index is sensitive to this:

$$\alpha = \frac{d\log(\lambda F_\lambda)}{d\log\lambda}\,, \tag{6.82}$$

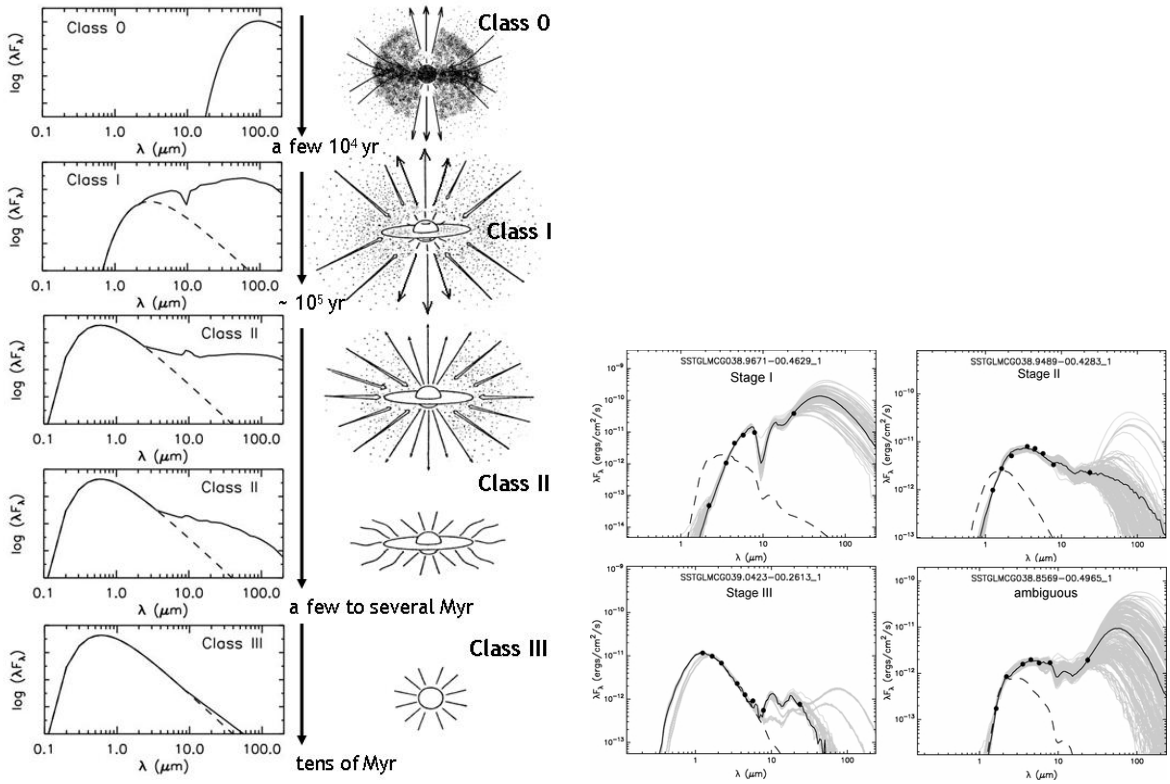


Figure 6.7: Star classes.

or $\nu F_\nu \propto \nu^{-\alpha}$. The classes are:
Class 0: objects are so heavily obscured that their spectra peak at wavelengths long-ward of 100 μm. Observationally, we typically do not even see Class 0 sources a mid-infrared wavelengths - we need far-infrared or sub-millimetric (peak) observations. Typical ages a few 10^4 yr.
Class I: More power radiated near 10 μm compared to 2 μm. This is the main accretion phase. Takes few 10^5 yr.
Class II: Still have accretion disks. This is the classic "T-Tauri" phase ("Herbig-Haro" objects are the higher mass versions). Takes few 10^6 years.
Class III: The accretion disks are now weak or absent. Takes a 10s of 10^6 years (faster for high mass stars).

A little terminology: The pre-main sequence (PMS) is before the main sequence! The zero-age main sequence (ZAMS) is when the star just reaches the main sequence, after the Class III phase.

The Schmitt-Kennicutt Law

Our Galaxy is making stars at a rate of about one Solar mass per year. The Galaxy is transforming molecular clouds into stars. If it were transforming all the molecular mass into stars, is would be making about 200 Solar masses of stars per year. It is not because of course most of the mass of GMCs is not undergoing gravitational collapse, and even the

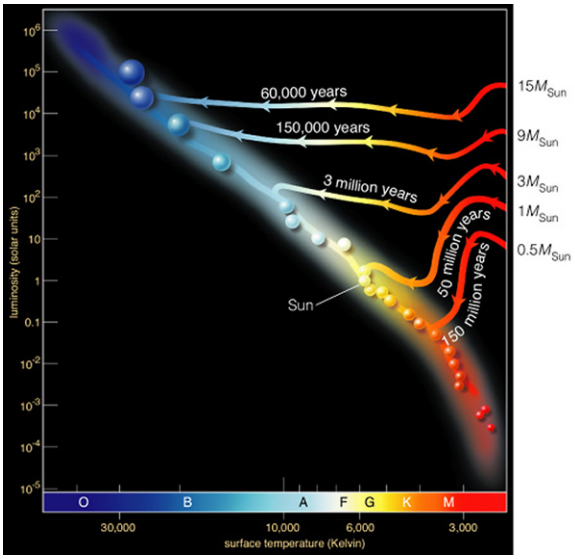


Figure 6.8: The stellar tracks on the H-R diagram for pre-main sequence stars.

mass that is collapsing is not all going in to a protostar - it is a very inefficient process.

Schmidt (1959) proposed that the star formation rate (SFR) should vary as a power of the local density ρ. Because the density is difficult to determine, Kennicutt (1998) proposed that the surface density star formation rate should vary as the surface mass density. Indeed it is! The Schmidt-Kennicutt law is remarkable in that it holds for low star formation and high star formation (starburst) galaxies.

The issue with the Schmidt-Kennicutt law is that we need some way to trace the star formation rate. All manner of methods have been used, from [NII] 205 μm emission, to free-free emission, to a census of stars, etc. Observationally, it has been found that the surface density star formation rate varies as the star formation rate to the 1.4th power:

$$\Sigma_{\rm SFR} = (2.5 \pm 0.7) \times 10^{-4} \left(\frac{\Sigma}{1\,{\rm M}_\odot {\rm pc}^{-2}} \right)^{1.4 \pm 0.15} \tag{6.83}$$

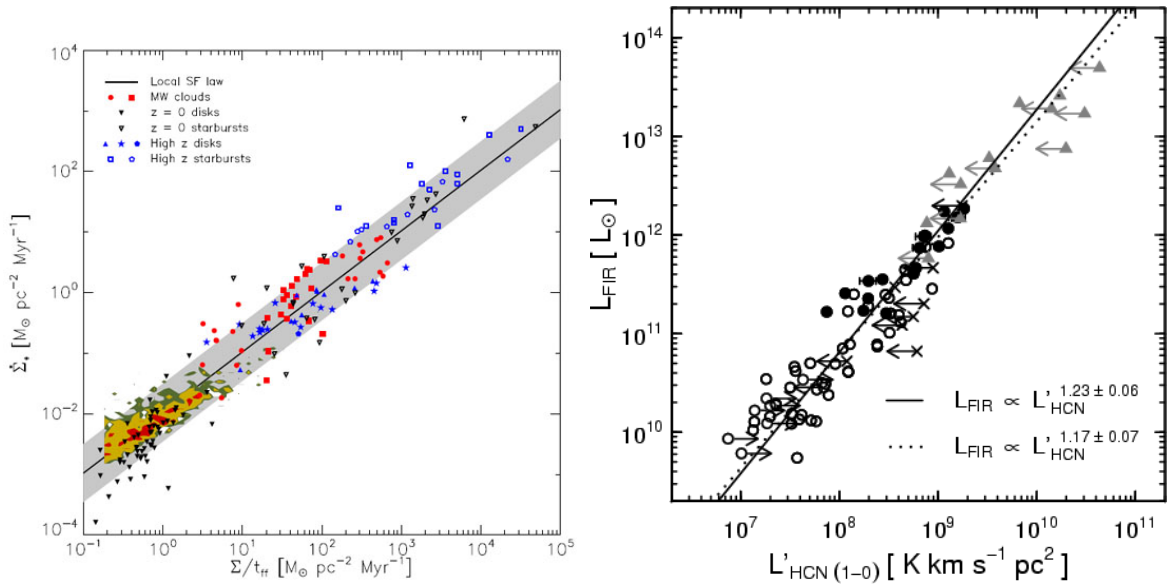


Figure 6.9: *Left:* Kennicutt-Schmidt relation for various galaxies. *Right:* Kennicutt-Schmidt-like relation looking at the far-infrared luminosity (a measure of star formation) versus the HCN luminosity (a measure of dense gas yet to form stars).